Estimation of Channelized Features in Geological Media Using Sparsity Constraint

by

Behnam Jafarpour

B.Sc. Civil Engineering
The University of Tehran, 1999

M.Sc. Civil and Environmental Engineering
The University of Delaware, 2003

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE IN PARTAIL FULLFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2008

Signature of Author .....
Department of Electrical Engineering and Computer Science
January 31st, 2008

Certified by .....
William T. Freeman
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by .....
Vivek K. Goyal
Assistant Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by .....
Terry P. Orlando
Chairman, Departmental Committee for Graduate Students

Reconstruction of Channelized Features in Geological Media Using Sparsity Constraint

by

Behnam Jafarpour

Submitted to the Department of Electrical Engineering and Computer Science
on January 31st, 2008 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Electrical Engineering and Computer Science

ABSTRACT

In this thesis, a new approach is studied for inverse modeling of ill-posed problems with spatially continuous parameters that exhibit sparseness in an incoherent basis (e.g. a Fourier basis). The solution is constrained to be sparse in the transform domain and the dimension of the search space is effectively reduced to a low frequency subspace to improve estimation efficiency. The solution subspace is spanned by a subset of a discrete cosine transform (DCT) basis containing low-frequency elements. The methodology is related to compressive sensing, which is a recently introduced paradigm for estimation and perfect reconstruction of sparse signals from partial linear observations in an incoherent basis. The sparsity constraint is applied in the DCT domain and reconstruction of unknown DCT coefficients is carried out through incorporation of point measurements and prior knowledge in the spatial domain. The approach appears to be generally applicable for estimating spatially distributed parameters that are approximately sparse in a transformed domain such as DCT. The suitability of the proposed inversion framework is demonstrated through synthetic examples in characterization of hydrocarbon reservoirs.

Thesis Supervisors:

William T. Freeman
Title: Professor of Electrical Engineering and Computer Science

Vivek K. Goyal
Title: Esther and Harold E. Edgerton Assistant Professor of Electrical Engineering

DEDICATION

To the precious memories of my brother PARVIZ.

ACKNOWLEDGEMENTS

I would like to express my gratitude to those who have had direct contribution to my research work in this thesis.

I wish to thank my thesis advisors, Professor Vivek K. Goyal and Professor William T. Freeman, for accepting to supervise my research work in this thesis and for their valuable advice and input in the past two years. I also would like to extend my gratitude to my PhD supervisor professor Dennis McLaughlin in the Department of Civil and Environmental Engineering for supporting me in pursuing my SM degree while completing my PhD work. I wish to thank Shell International Exploration and Production for their financial support. I also would like to acknowledge Schlumberger for donating ECLIPSE reservoir simulator that was used in this research. I appreciate the assistance from the staff at EECS and CEE departments. In particular, I am thankful to Janet Fisher, Kris Kipp, and Jeannette Marchockie for their help with administrative issues.

I have been privileged to meet and make friends with outstanding individuals who made my years in Boston and MIT unforgettable and precious. It is not possible to mention everyone's name in here; however I would like express my appreciation for their friendship and support and for all the things I have learned from them. I am grateful for the support, inspiration, and understanding that I have received from my girlfriend, Maryam Modir-Shanechi, during the completion of this thesis. The past few years would not have been as memorable, productive, and beautiful without her in my life.

Above all, I would have never been where I am if it were not for my family's unconditional love, support, sacrifice, and patience. My utmost respect and heart-felt gratitude go to my mother, Effat Gholizadeh Zargar, who has been a great example of selflessness to me. I am grateful to my father, Ahmad, for every thing he has done for me. I am thankful to my brothers Aref, Parviz (may his soul rest in peace), Behrouz, and my sister Fatemeh for supporting me at different stages of my life and during this long journey away from home. I bring this section to end by expressing my special thanks to my Brother Behrouz for his support and encouragements.

Contents

1. Introduction and Problem Statement

2. Methodology and Problem Formulation

3. Results and Discussion

4. Inversion Using Dynamic Observations

5. Conclusions and Future Directions

List of Figures

Chapter 1

Introduction and Problem Statement

1.1 Introduction and Significance

The inverse problem of estimating patterns and structures using uncertain prior models and limited point measurements is encountered in several engineering and science applications, from geophysics to medical imaging for diagnosis. These problems are known to be severely ill-posed and challenging to solve. For example, estimating subsurface structures/patterns such as channels and faults using discretized pixel-based descriptions of them in the spatial domain is futile unless important prior assumptions about these structures are built into the solution algorithm. These assumptions limit the application of these algorithms when prior knowledge is not available or is inaccurate. A similar issue is encountered in pattern recognition, computer vision, and medical applications in which limited point data sets are used to infer the shape of an existing feature.

The commonality between these problems lies in the presence of a structure (pattern) in the solution. Patterns are nearly constant (continuous) features that can have irregular shapes, which makes their parametric description in spatial domain very challenging. Therefore, it is a common practice to resort to pixel-based descriptions that can introduce significant redundancy. The two-dimensional discrete cosine transform (DCT) efficiently approximates image data by encoding the main information in continuous features into a few coefficients in the Fourier domain. Patterns and structures lend themselves to such parsimonious descriptions due to their continuity in the spatial domain. In this thesis, a compressed sensing framework is proposed for reconstruction of structural patterns in a compression domain (such as DCT) by using limited point observations of the original feature in the spatial domain and constraining the solution to be sparse in the DCT domain. The sparsity assumption is the key behind the new formulation and stems from the continuity that is ubiquitous in most of these applications. For instance, the prevailing geological continuity of the subsurface channels in the spatial domain translates into nearly sparse representation in the DCT domain, which fits well into the compressive sensing problem formulation proposed here.

In the proposed approach of this thesis, patterns and structures are estimated in a compression domain (such as Fourier or discrete cosine transform) and the inverse problem is defined using the compressive sensing framework to arrive at an efficient and better-posed algorithm. Compressive sensing has recently been a topic of great interest in statistics and signal processing. While theoretical aspects of this method are currently being investigated, its immediate application in other areas such as geophysics, pattern

recognition, and medical imaging needs to be studied. Compressive sensing theory uses assumptions such as randomness of the observation points and strictly sparse signals that may seem to limit its application in realistic systems. Relaxing some of these assumptions to some extent (to make the approach more practical) may not affect the quality of the estimation results dramatically. The approach proposed here is evaluated using examples from subsurface characterization and identification of geological channels in hydrocarbon reservoirs.

1.2 Compression and Inverse Problems

The inverse problem of estimating high-dimensional spatially distributed parameters from limited point measurements can be better posed through their low-rank representation. This reduces the dimension of the underlying model parameters by providing an effective description in a suitable coordinate system and eliminating parameters that represent insignificant (high frequency) details. This approach is often taken in image compression [1,2].

Compression transform algorithms that are used for parsimonious signal representation may also be applied to parameterization of ill-posed problems. A specific example of this is in characterization of hydrocarbon reservoirs. Petrophysical reservoir properties, such as porosity and permeability, are highly correlated spatial parameters that determine fluid flow displacement within rocks and oil production behavior, which is of great interest to reservoir engineers. Reservoir simulators use these parameters as inputs to solve a set of PDEs that represent flow movement in porous media. In practice, these model parameters

are far from known; as a result model predicted production patterns are highly uncertain. Identification of these input parameters using available knowledge and observations is essential for accurate characterization of subsurface media and optimal development of the resources they contain.

The commonality between image compression and parameterization lies in the sparse basis used to describe the underlying image. However, for a pre-specified basis and a known image, the significant (energy carrying) basis vectors are uniquely identified by coefficients representing the transformed image. The image and its transform are both *known* and compression is achieved by truncating insignificant basis elements. In contrast, parameterization is usually done to constrain the solution space in an inverse problem and to avoid redundant computation in the original parameter space that may lead to unrealistic estimation results. In this case, the solution (parameter field) and its transform are *unknown* a priori. However, prior knowledge and observations may be among available sources of information that can be used for identifying the appropriate basis vectors and estimating their corresponding coefficients.

In this thesis a few examples are used to show how these information sources can be combined with a reduced representation to provide an efficient estimation framework. A compressed sensing approach is then proposed to improve the estimation performance. The particular inverse problem that has been considered in the examples of this thesis is drawn from characterization of petroleum reservoirs using static and dynamic point

measurements of their attributes and states. To provide context, the inverse problem of interest is described first.

1.3 Mathematical Modeling of Fluid Flow in Porous Media

Reservoir engineers use a mathematical model to represent, analyze, and quantify a reservoir's flow displacement pattern and production behavior. Predictions of future performance can be used to prepare sound reservoir development and management strategies. Therefore, reservoir modeling plays a central role in planning oilfield operational activities.

The general form of the governing equations for the two phase (oil-water) immiscible flow in porous media is derived from mass and momentum conservation principles [3,4]:

$$\nabla.\left(\frac{\lambda_o}{B_o}\mathbf{k}(\nabla P_o - \gamma_o \nabla Z)\right) = \frac{\partial}{\partial t}(\phi \frac{S_o}{B_o}) + q_o \tag{1.1}$$

$$\nabla.\left(\frac{\lambda_w}{B_w}\mathbf{k}(\nabla P_w - \gamma_w \nabla Z)\right) = \frac{\partial}{\partial t}(\phi \frac{S_w}{B_w}) + q_w \tag{1.2}$$

Here, λ_o and λ_w represent mobility of oil and water (measure of the ease with which a fluid can be moved in a specific medium), B_o and B_w are the formation volume factors (volume of fluid as function of pressure relative to its volume at standard pressure), $\boldsymbol{k}$ is intrinsic permeability (a physical property of rocks that indicates their conductivity), P_o and P_w are oil and water pressures, γ_o and γ_w represent oil and water densities, Z is elevation, S_o and S_w are oil and water saturation (ratio of oil (water) volume to total

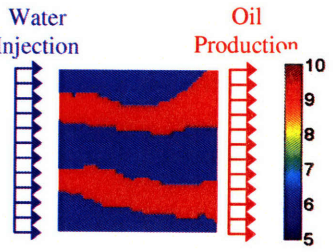
volume of pore space in rocks), and finally q_o and q_w are sink and source (control) terms referring to injection and production rates per unit volume.

These two equations have four unknown dynamic variables, namely P_o, P_w, S_o, S_w. For a given set of model input parameters, boundary conditions, initial conditions, and reservoir controls (well rates and/or pressures) two additional (*constitutive*) equations are used to find a *unique* solution for these unknowns at any given time. The constitutive equations reflect capillary pressure (pressure difference between the two phases at any given point) relations and physical saturation constraint for a given control volume:

$$\begin{aligned} P_o - P_w &= P_c(S_w) \\ S_w + S_o &= 1 \end{aligned} \tag{1.3}$$

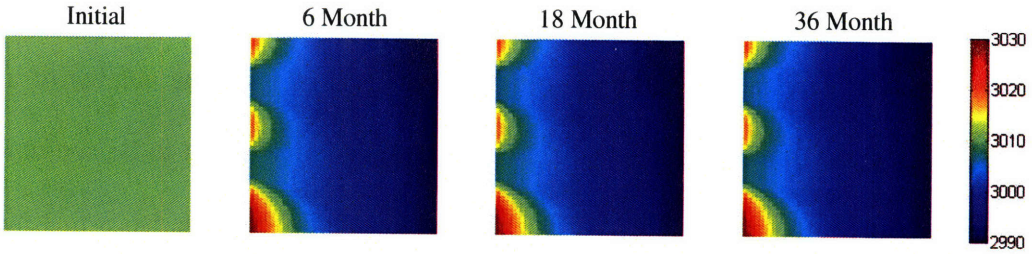
Forward integration of equations (1.1)-(1.3) provides model solutions in time that are used to predict flow behavior within the reservoir. In general, these equations need to be discretized and numerically solved to obtain reservoir states (saturation and pressure) in time. Figure 1.1 shows the solution of these equations in a waterflooding experiment for a given set of input parameters. Waterflooding is a secondary drive production method (after natural depletion), in which water is injected to push the resident oil toward the production wells and to maintain high reservoir pressure. It is important to observe that the shape of the high permeability channels (shown in red) strongly affects saturation and pressure profiles (Figure 1.1).

a) Permeability and injection/production scenario



2D Reservoir

b) Pressure solution



c) Saturation solution

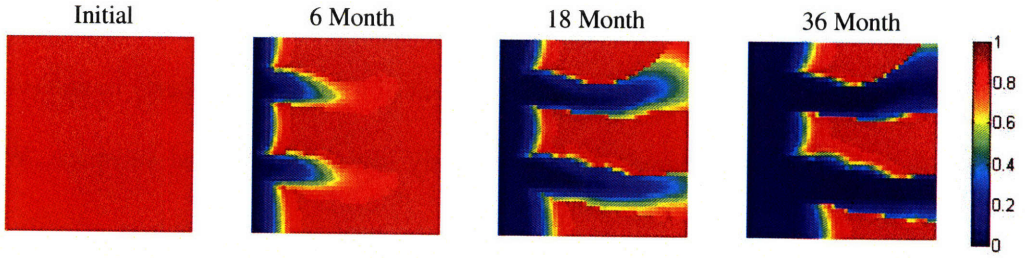


Figure 1.1 Simple waterflooding experiment to demonstrate solution to the multiphase flow equations in porous media (equations (1)-(3): a) permeability field and horizontal injection production wells (64 injection wells (left) and 64 production wells (right)); b) pressure field solution in time; c) saturation profile solution in time.

1.4 Reservoir Inverse Modeling

Solution of the PDE equations (1.1)-(1.3) is only useful when accurate model inputs are available assuming the model adequately captures the physics of the system. In practice, however, several sources of error exist that make the solution of these equations uncertain. It is, therefore, important to calibrate reservoir models by tuning reservoir parameters to match reservoirs observed past performance (a process referred to as *history matching* in the reservoir engineering literature).

When pixel-based description is adopted, history matching of even and up-scaled model of a heterogeneous hydrocarbon reservoir is an extremely underdetermined problem, i.e. the number of unknown parameters is significantly larger than the number of available data. As a result, the solution of the estimation problem is known to be non-unique [5]. This results in several solutions with different geological characteristics that match the observed past data equally well but provide incorrect forecasts.

The inherent geological continuity (correlations) in a reservoir makes pixel-based descriptions inefficient. The information content of a reservoir property image (volume in 3D cases) can be represented in a significantly lower dimensional subspace. While the examples used in this thesis are simplified, the proposed approach is effective due to strong spatial correlation that is dominant in many realistic reservoirs (per geological continuity).

The major focus of this thesis is on improving the stated ill-posed problem by reducing the size of the unknown parameters while preserving their important features. To do this, a widely used sparse representation basis, discrete cosine transform (DCT) is used [6,7]. While DCT bases are used for compressing known images in the image compression literature [6,7], in inverse modeling it can also be used for sparse representation of unknown parameters (i.e. parameterization). This distinction makes identification of significant basis vectors an interesting problem in inverse modeling.

Since permeability is a major source of uncertainty in reservoir modeling and plays a prominent role in governing flow displacements (see Figure 1.1), it is considered as the only unknown to be estimated here. Furthermore, the estimation approaches in this thesis require the solutions of the flow equations (state-space model), in (1.1)-(1.3), which are obtained using a commercial reservoir simulator [8].

1.5 Prior Information

Reasonable probabilistic models for the permeability field can be constructed from seismic and geological surveys. In this thesis, unconditional permeability realizations were generated using a channelized training image shown in Figure 1.2. This training image has 250 × 250 × 1 grid blocks comprising two rock (litho-facies) types: low permeability background (blue) and high permeability channels (red). The training image has mainly horizontal (east-west) channels that are believed to be present at the site of interest.

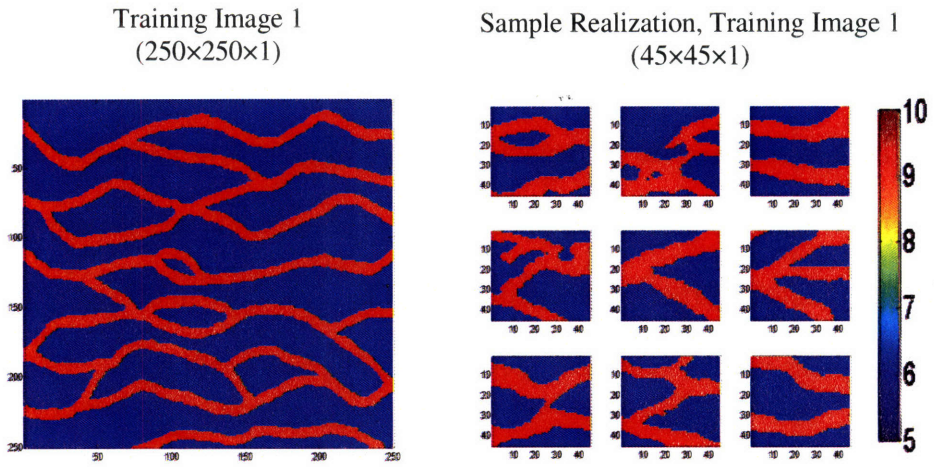


Figure 1.2 Permeability training image (left) and nine sample permeability realizations (right), generated by SGeMS using multiple point geostatistics.

The permeability realizations in Figure 1.2 are generated using the multiple-point geostatistical algorithm *SNESIM* [9,10]. Each realization is discretized over a 45 × 45 × 1 (450m × 450m × 10m) grid block system. The varying shape and geometry of the channel facies in these realizations are the major source of uncertainty. The highly structured nature of the facies distribution in these realizations suggests a distinctive preferential flow displacement pattern. In addition, a high level of correlation (redundancy) is observed in description of these facies, suggesting that the field could be represented more efficiently if a more appropriate description were adopted.

1.6 Sparse Representation through Linear Transformation

A general unitary transformation of a one dimensional sequence $\{u(n),\ 0 \le n \le N\text{-}1\}$ can be expressed as convolution of $u(n)$ with a specified function $a(k,n)$ [1,2]:

$$v(k) = < u(n), a(k,n) > = \sum_{n=0}^{N-1} a(k,n).u(n) \quad \text{for} \quad 0 \le k \le N-1 \qquad (1.4)$$

The original sequence can be reconstructed by applying the inverse transform $a^*(k,n)$ to the transform coefficients:

$$u(n) = < a^*(k,n), v(k) > = \sum_{k=0}^{N-1} v(k).a^*(k,n) \qquad \text{for} \quad 0 \le n \le N-1 \qquad (1.5)$$

where $\{a^*(k,n),\ 0 \le n \le N\text{-}1\}^T$ is the inverse transform kernel. It is often possible to construct a good approximation to $u(n)$ with a truncated version of the inverse transform. In particular, we compute and retain only the first $K_r << N$ expansion terms of $u(n)$ if the remaining $N\text{-}K_r$ terms have small contribution in the summation:

$$v(k) = \sum_{n=0}^{N} a(k,n).u(n) \qquad \text{for} \qquad 0 \le k \le K_r - 1 \qquad (1.6)$$

$$u(n) \approx \sum_{k=0}^{K_r-1} v(k).a^*(k,n) \qquad \text{for} \qquad 0 \le n \le N-1 \qquad (1.7)$$

For a unitary transform, the terms with small coefficients are omitted from the $u(n)$ expansion since they make a small contribution to the signal energy. In this case, the retained coefficients in the expansion represent a compressed version of the original signal [1,2]. The truncated basis coefficients vector $v(k)$ provides a sparse version of $u(n)$ that requires less transmission time and storage than the original *known* image [1,2].

1.7 Discrete Cosine Transform

The discrete cosine transform (DCT) is a linear transform that is widely used for image compression due to the sparsity of its basis vectors. The one dimensional forward DCT of a signal $u(n)$ of length N has the following form [6,7]:

$$v(k) = \alpha(k)\sum_{n=0}^{N-1} u(n).\cos\left[\frac{\pi(2n+1)k}{2N}\right] \quad 0 \le k \le N-1 \tag{1.8}$$

where $\alpha(k)$ is defined as:

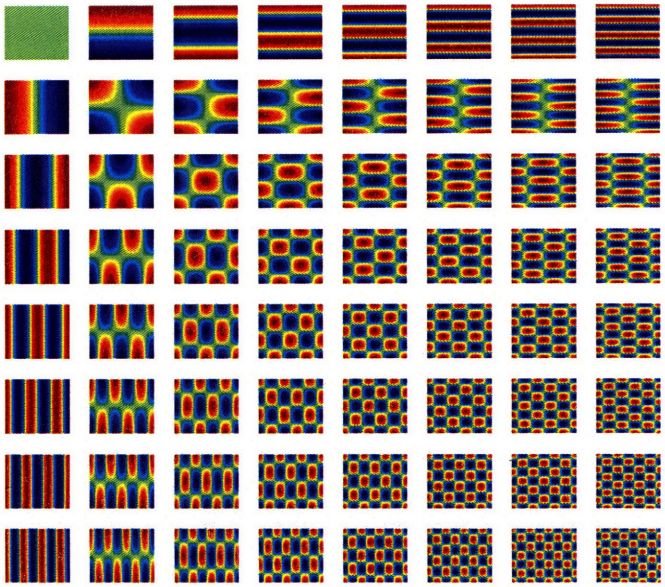
$$\alpha(k) \equiv \begin{cases} \sqrt{\frac{2}{N}} & k=0 \\ \sqrt{\frac{1}{N}} & 1 \le k \le N-1 \end{cases} \tag{1.9}$$

The inverse DCT can then be written as :

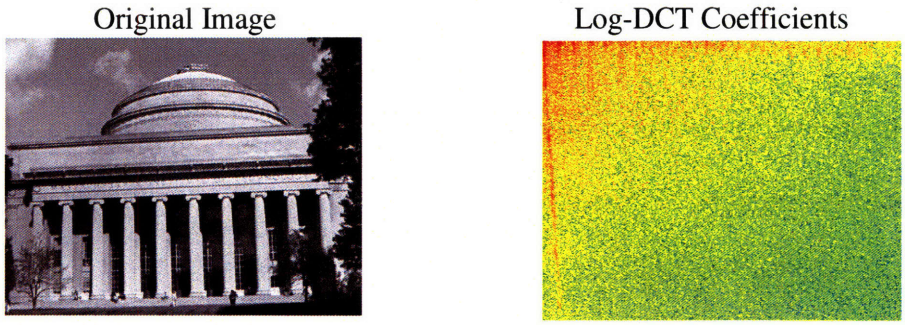
$$u(n) = \sum_{k=0}^{N-1} \alpha(k)v(k).\cos\left[\frac{\pi(2n+1)k}{2N}\right] \quad 0 \le n \le N-1 \tag{1.10}$$

Extension of the above equations to higher dimensional signals (images and volumes) is straightforward [7,11]. However, the separability property of DCT bases can be exploited to achieve computational savings by applying the one dimensional transform in each direction [7,11]. Figure 1.3a shows sample low frequency image bases that can be used for representing 45-by-45 images. The basis images are arranged according to their orientation and level of detail in a descending order from upper left to lower right. Depending on the desired level of details in the approximation more high frequency components (lower right basis images) are included.

(a) Low-frequency discrete Cosine transform modes



(b) An example image and its log-DCT coefficients



(c) Low-rank approximations with increasing number of DCT modes

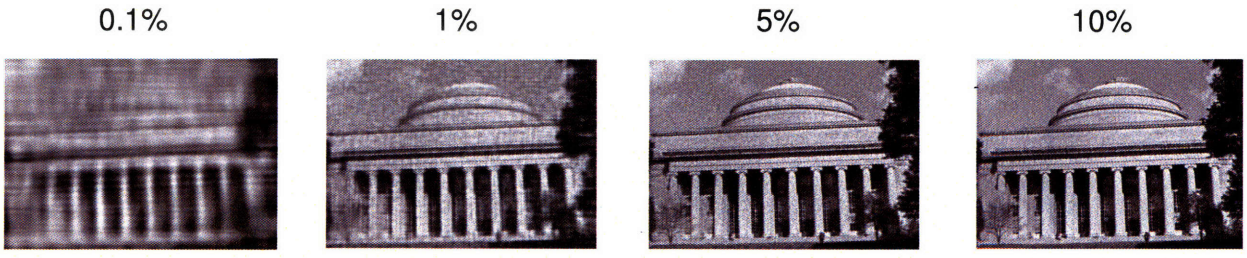


Figure 1.3 Compression power of the discrete cosine transform: (a) sample discrete cosine transform modes (bases); (b) an example image (the famous MIT dome) with its log-DCT coefficients; (c) approximate representations of the example image with increasing number of included modes.

Figure 1.3b shows an image of the MIT dome with a truck on top of it (first column) and the DCT coefficients for this image, using the same ordering convention as in Figure 1.3a (second column). Figure 1.3c illustrates how the DCT can be used to compress the original image. It shows the approximation using the largest 0.1%, 1%, 5%, and 10% of the original DCT coefficients. The concentration of the large coefficients on the top left corner (low frequency modes) of the DCT coefficients in (middle right panel of) Figure 1.3b is apparent. This clustering of coefficients generally corresponds to the modes with large scale variations in the horizontal, vertical, and diagonal directions. It is clearly seen in Figure 1.3c that after including only 5% of the DCT coefficients most of the details in the original image are retrieved.

1.8 Summary of Problem Statement

The introduction presented in previous sections covered several important components that can be combined to formulate the problem. The inverse problem of interest is summarized in this section.

We are given the following knowledge about a synthetic test reservoir:

i. Equations (1.1)-(1.3) as the governing equations describing multiphase flow in porous media.

ii. All inputs into equations (1.1)-(1.3) except the permeability **k**, which is the unknown parameter.

iii. The permeability field **k** is (nearly) sparse in a predefined transform domain (DCT in this case).

iv. Point observations of the permeability field at well locations.

v. Prior knowledge of the permeability field in the form of a training-image or training library that portray structurally similar features to the true permeability.

vi. Observations of reservoir production variables at well locations (flow rates and pressures at well locations).

An efficient estimation technique is desired that integrates all these sources of information to provide an accurate estimate of the unknown permeability field.

This problem is formulated and solved in the later chapters of this thesis. Chapters 2 and 3 consider solution of the problem without dynamic observations (i.e. an interpolation problem). Chapter 4 presents the solution when dynamic sources of information are also available for integration. Finally, Chapter 5 summarizes the conclusions of this work and discusses possible future research directions.

Chapter 2

Methodology and Problem Formulation

2.1 Basic Compressed Sensing Formulation

Compressed sensing [12-15] is a recently introduced paradigm for estimation or perfect reconstruction of sparse signals from partial observations in a complementary, "incoherent" domain using convex optimization. It has attracted researchers' attention in several disciplines including signal processing and statistics. A simple formulation of the approach is given in this section. A more general formulation is presented later in the chapter while further mathematical details are left to original publications on this topic [12-15].

Assume a sparse signal x_N with sparsity S (a signal with S non-zero coefficients) and its transformation coefficients y_N under the transformation matrix $\Phi_{N\times N}$:

$$y_{N\times 1} = \Phi_{N\times N}\, x_{N\times 1} \tag{2.1}$$

Reconstruction of the signal x_N using only K<<N observations of it in the transformed domain (y_K):

$$y_k = \langle x, \varphi_k \rangle, \quad k = 1,....K \quad \text{or} \quad y = \mathbf{\Phi} \cdot x \tag{2.2}$$

is desired, where x ϵR^N , φ ϵR^N, and y ϵR^K:

$$y_{K\times 1} = \mathbf{\Phi}_{K\times N}\, x_{N\times 1} \tag{2.3}$$

The reconstruction is shown [13,14] to be perfect under specific conditions by solving the following l_1 norm constrained minimization:

$$\min_{x\in R^N} \| \tilde{x} \|_1 \quad \text{subject to} \quad y_{K\times 1} = \mathbf{\Phi}_{K\times N}\, \tilde{x}_{N\times 1} \tag{2.4}$$

This minimization problem can be posed [16] as a linear program of the form:

$$\min_{z\in R^N} 1^T \cdot z \quad \text{subject to} \quad y_{K\times 1} = \mathbf{D}_{K\times 2N}\, u_{2N\times 1} \tag{2.5}$$

where $u_{2N\times 1} = \begin{bmatrix} u^+ & u^- \end{bmatrix}^T$ with $\tilde{x} = u^+ - u^-$ and $D_{K\times 2N} = \begin{bmatrix} \mathbf{\Phi}_{K\times N} & -\mathbf{\Phi}_{K\times N} \end{bmatrix}$. For an appropriate choice of the incoherent basis $\mathbf{\Phi}$ (one in which the signal is not sparse), and a signal x with sparsity S, reconstruction is exact with very high probability when $K \geq C \cdot S \cdot \log N$ [13,14], where C is a constant between 3 to 5 [13,14]. A critical assumption for this result to hold is sparsity. It is natural to ask "*Is it reasonable to assume that realistic reservoir properties such as permeability are sparse in DCT (or any other transform) domain?*" In most cases the answer is yes, at least as a first order approximation. This is due to the geological continuity of rock formations that often exhibit strong correlations. The decorrelating power of DCT basis can be exploited to provide sparse representation

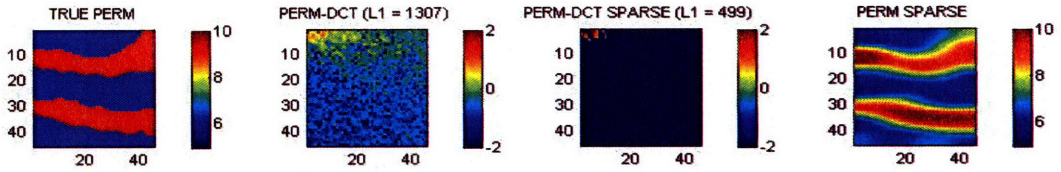
(approximation) of reservoir properties in the DCT domain. This is shown with an example in the next section.

2.2 A Simple Interpolation Example

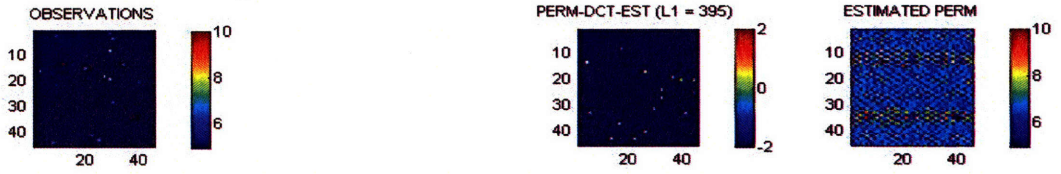
An example is used in this section to illustrate the above formulation in the context of reservoir characterization. The DCT basis is used as transformation matrix (Φ). Figure 2.1a shows a sample permeability field and its corresponding DCT coefficients magnitude (in logarithmic scale) after transformation (first and second columns respectively). Although the coefficients are not exactly zero, most of them are small and can be zeroed out without a major loss in quality (third column). The third and fourth columns in Figure 2.1a show the (S=15) largest DCT coefficients and their corresponding approximate representation, respectively.

For K observed pixels of this permeability in spatial domain, reconstruction of the DCT coefficients was carried out following the above formulation. The results for K = 20, 40, 60, 80, 100 are shown in Figures 2.1b-2.1f. After including 100 random observations the original signal was perfectly constructed in almost all trials with different sets of random observations. Examination of Figure 2.1 indicates that perfect reconstruction is not possible for smaller number of observations due to existence of a solution with smaller l_1 norm that matches the observations perfectly. In other words, the observations do not constrain the solution sufficiently. As the number of observations increases it becomes less likely to fit the observations with a set of DCT coefficients that have a smaller l_1 norm than the true solution. Since the problem has a unique global minimum (due to

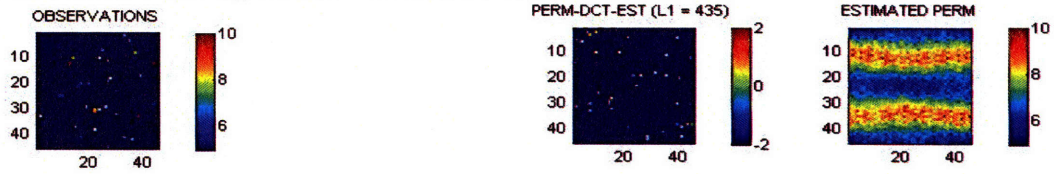
a) True Permeability Field (Spatial and DCT domains)



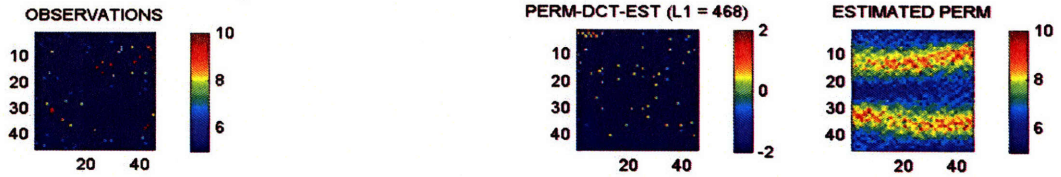
b) Estimated Permeability with K = 20 observations.



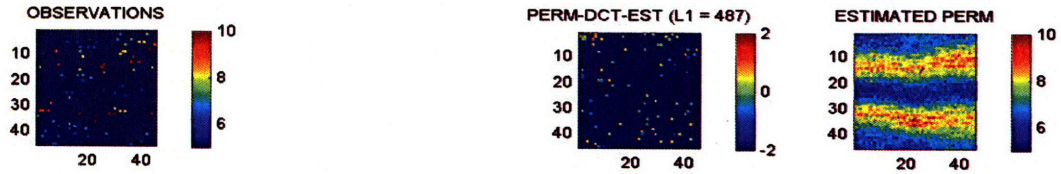
c) Estimated Permeability with K = 40 observations.



d) Estimated Permeability with K = 60 observations.



e) Estimated Permeability with K = 80 observations.



f) Estimated Permeability with K = 100 observations.

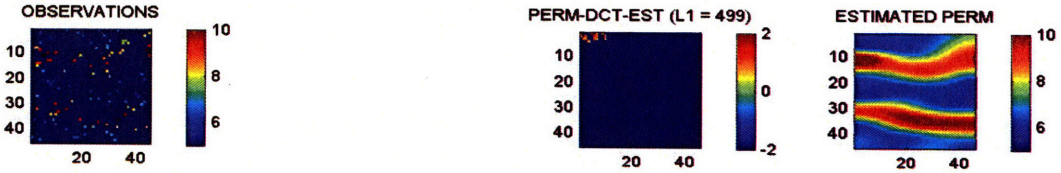


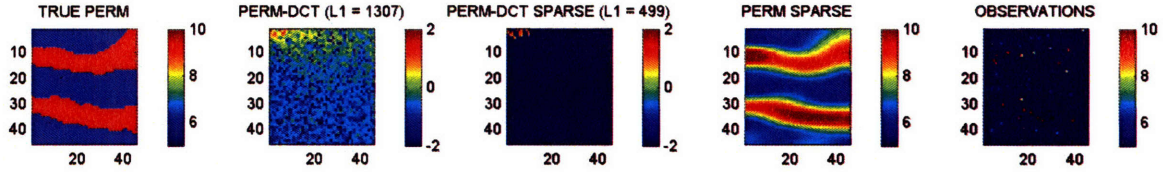
Figure 2.1 Interpolation example using compressed sensing formulation: True and DCT representation of the permeability (full and approximated) are shown in (a). Rows b-f show the reconstructed permeability using K = 20, 40, 60, 80, and 100 random observations (observations (first column), estimated DCT coefficients (third column), and the estimated spatial permeability field (last column) are shown).

convexity), in the limit when $K \rightarrow C \cdot S \cdot \log N$ the minimum l_1 solution is expected to converge to perfect reconstruction [13,14]

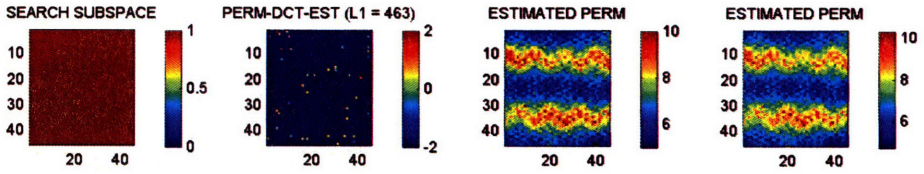
Another important observation is the distribution of the DCT solution coefficients when perfect reconstruction is not achieved. The wide frequency spread in the coefficients suggest that by imposing certain constraints on the solution, such as excluding high frequency modes, the probability of achieving perfect reconstruction can be increased. This is demonstrated in Figure 2.2. In this figure, K=40 observations are used to constrain the reconstruction. The experiment is run using three different scenarios. Figure 2.2b shows the representation when the entire N=2025 space is searched to find the correct feature in sparse permeability field. It is seen that some of the estimated DCT coefficients are selected from the high-frequency region. Figure 2.2c and 2.2d show the same experiment when N = 210 and N = 120 low-frequency bases are included in the search. It is clear that reducing the search space results in a solution that is more representative in this example. This has important implications for reservoir characterization where prior knowledge of the permeability field is available and can be used to constrain the search subspace.

However, for perfect reconstruction, the number of observations is proportional to the logarithm of the search dimension. This implies that exponential reduction in dimension of search space (N) is needed to reduce the number of observations in a linear manner. This is important in reservoir characterization where only limited spatial observations are available (small K) due to the cost associated with data acquisition. Therefore, it may be

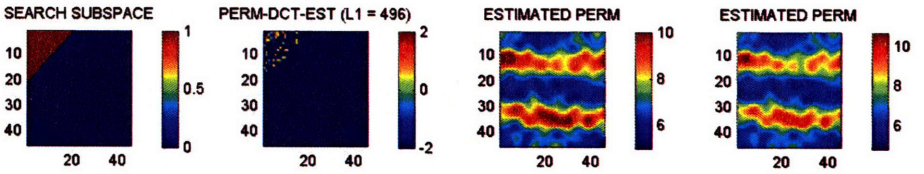
(a) True Permeability with its sparse (S = 15) representation and K = 40 observations



(b) Reconstructed permeability with N = 2025 dimensional search space



(c) Reconstructed permeability with N = 210 dimensional search space



(d) Reconstructed permeability with N = 120 dimensional search space

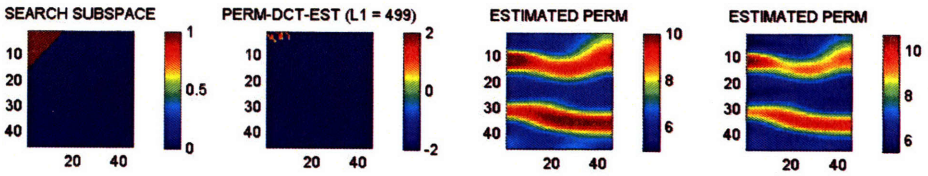


Figure 2.2 Compressed sensing reconstruction using a reduced low-frequency subspace: (a) true permeability and its DCT representation (first and second columns) and their corresponding sparse (S=15) representation (third and fourth columns) are shown as well as K = 40 observations (fifth column); (b) Reconstruction of the permeability shown in (a) using the original DCT space N = 2025); (c) Reconstruction of the permeability shown in (a) using a reduced low-frequency subspace of dimension N = 210; (c) Reconstruction of the permeability shown in (a) using a reduced low-frequency subspace of dimension N = 120. The search spaces in (b)-(d) are shown with the masks plot (triangle).

possible that sufficient measurements are not available to constrain even a reduced representation of the signal.

2.3 A More General Formulation

The preceding examples suggest that the sparsity of the parameter of interest can be exploited to develop a more efficient and better posed estimation scheme. While above results are encouraging, they are overly simplified and further development of the approach is required to solve a full-blown dynamic inverse problem. A few generalization of this formulation is considered in this section.

When reliable prior knowledge of the unknown parameters is available the inverse problem is further constrained and the reconstruction algorithm is improved. In addition, specifying equality (hard) constraint for observations can be too restrictive and in practice observation errors need to be taken into account. This can be achieved by including the observations as soft constraints. Further, while in the compressed sensing formulation the l_1 norm is typically used for reconstruction of sparse signals, a generalization can be considered in which other norms are used in the problem formulation. A few alternative norms will be considered in the next section.

A more general formulation of the above problem that incorporates prior information and allows for observation error can be written as:

$$\min_{x \in R^N} \| C_y^{-1/2} (y - \Phi \tilde{x}) \|_p + \gamma \| W \cdot \tilde{x} \|_q \qquad (2.6)$$

The first term in the objective function represents the penalty for deviation from the observations. The matrix $C_y^{-1/2}$ contains the information about the quality of each observation. The second term is often referred to as regularization and is used to improve the ill-posedness of the inversion. In our application, **W** is the weighting matrix that can be computed using the prior information (training images). The weighting coefficients can be used to include (exclude) relevant (irrelevant) basis vectors in the approximation and give more weight to the ones that are likely to have significant contribution to approximation of the unknown permeability field. The coefficient γ in front of the second term is used to adjust the fidelity to prior versus observations.

2.3.1 Basis Training Procedure

A training procedure can be developed to obtain the weighting coefficients (elements of **W**) through the prior permeability library. This is briefly described in this section.

When a prior library (or training image) exists that contains features that are believed to be structurally similar to the unknown image, it can be utilized to determine the significance of each basis vector in the reconstruction. The weighing matrix **W** is computed using the following procedure:

1) The DCT transform of each image in the library is computed
2) The magnitude (absolute value) of the DCT coefficients is averaged across the library

3) The resulting mean of magnitudes is truncated by specifying a threshold to remove basis vectors that are likely to have insignificant contribution (small magnitude).

4) The remaining coefficients are inverted and normalized to yield **W**.

The matrix **W** obtained from the above procedure ensures that in the reconstruction algorithm significant basis vectors (as suggested by prior library) are given small penalties so they can assume large values, whereas large penalties are associated with less significant basis vectors to keep their corresponding coefficients small.

Figure 2.3 illustrates the training procedure used to obtain **W**. In this figure, a threshold specified to retain 78 coefficients in the expansion.

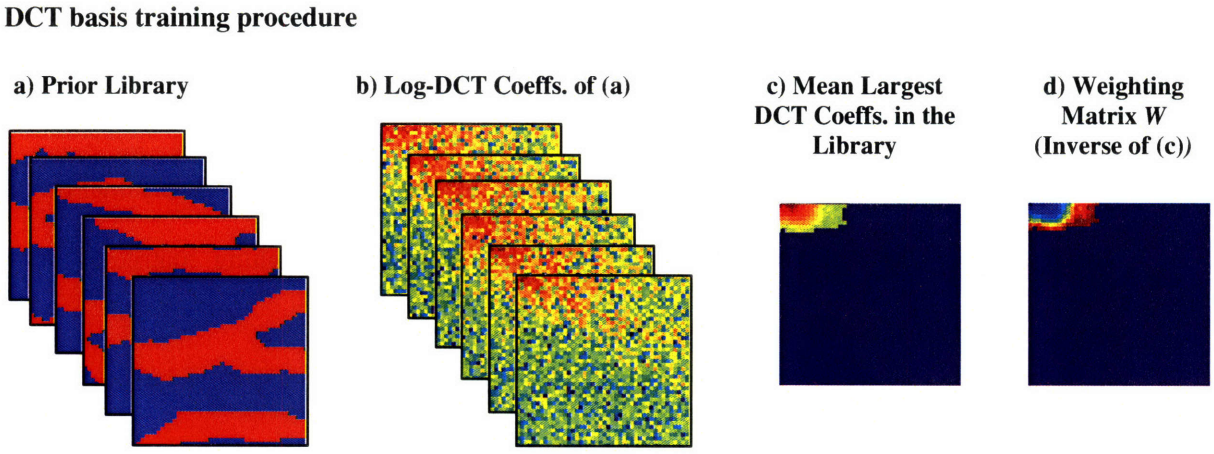


Figure 2.3 Basis training procedure using a library (ensemble) of prior models: (a) prior images used in the training; (b) the logarithm of DCT coefficients magnitudes; (c) the largest (over prior library) 78 DCT coefficient magnitudes selected to represent the important DCT basis vectors for reconstruction; (d) the weighting matrix obtained by taking the inverse of the DCT coefficients in (c).

It is important to note that if the prior image library does not represent the unknown features accurately, the reconstruction results will be adversely affected. Therefore, when prior is specified the reconstruction outcome is less robust and is sensitive to prior information [17].

2.4 Solution Approaches

This section discusses the solution to the general formulation presented above. Based on the specified norms l_q and l_p, three alternatives objective functions are considered. These formulations can be derived from the Bayesian approach to regularization with either Gaussian or Laplacian distribution assumptions for observation noise and/or prior statistics [16,18,19,20]. When these statistical assumptions are applicable, the solutions to the formulations presented in the following sections provide the *Maximum A Posteriori (MAP)* estimate.

Several studies have been conducted on the distribution of DCT coefficients for various types of images including natural and medical images. Different probabilistic distributions have been proposed to model the transform coefficients [21-24]. Some studies have found the Laplacian distribution to be a better fit for the DCT coefficients while others have reported that no single distribution can be used to model individual DCT coefficients [22]. Recently, the Generalized Gaussian Function (GGF), which includes the Uniform, Laplacian, and Gaussian distributions as special cases, has been used to model the distribution of the DCT coefficients [24]. The lack of agreement in

these studies suggests that DCT coefficients distribution can vary depending on the image/data sources and types.

In the following formulations, both Laplace and Gaussian distributions are considered as prior statistics for the DCT coefficients. However, it is generally agreed that l_1 norm is a better choice for representing sparsity. This will be studied in the examples of next chapter.

2.4.1 Linear Least Squares (*LLS*) Solution ($q = p = 2$)

A formulation that is obtained by assuming Gaussian additive noise for observations and a regularization term [25] to represent Gaussian prior statistics takes the following form:

$$\min_{x \in R^N} \| C_y^{-1/2} (y - \Phi \tilde{x}) \|_2 + \gamma \| \mathbf{W} \cdot \tilde{x} \|_2 \qquad (2.7)$$

The second term in the *LLS* formulation is often used to enforce regularity (e.g. when $\mathbf{W}$ is a difference operator or inverse of the squared root of covariance). Here, the l_2 norm is used to compare its performance with the more widely used l_1 norm sparsity constraint. If the observations have additive white noise with Gaussian distribution, and the prior is also Gaussian with exactly specified covariance, the *LLS* solution can be interpreted as the "maximum a posteriori" (MAP) estimator [19,25].

A closed form solution of the above minimization can be obtain by setting the derivative of the objective function with respect to $\tilde{x}$ equal to zero. This results in the following closed-form solution:

$$\tilde{x} = \left[\gamma^2(\mathbf{W}^T\mathbf{W} + \mathbf{W}\mathbf{W}^T) + \mathbf{\Phi}^T(C_y^{-T/2}C_y^{-1/2} + C_y^{-1/2}C_y^{-T/2})\mathbf{\Phi}\right]^{-1}\mathbf{\Phi}^T\,(C_y^{-T/2}C_y^{-1/2} + C_y^{-1/2}C_y^{-T/2})\,\mathrm{y} \tag{2.8}$$

For symmetric $\mathbf{W}$ and $C_y^{-1/2}$ this will reduce to:

$$\tilde{x} = \left[\gamma^2 \cdot \mathbf{W}^T\mathbf{W} + \mathbf{\Phi}^T C_y^{-1}\mathbf{\Phi}\right]^{-1}\mathbf{\Phi}^T C_y^{-1}\,\mathrm{y} \tag{2.9}$$

This equation is widely known as the *LLS* solution.

2.4.2 Least Absolute Deviation (*LAD*) Solution ($q = p = 1$)

If the above Gaussian assumptions are not applicable, the *LLS* formulation may not give a good representation. For instance, in image processing due to presence of edges the Gaussian prior assumption is often violated. In addition, the l_1 norm is well known for facilitating automatic order selection [16,26], which is desirable in this application. Furthermore, in some cases the additive noise may follow a Laplace distribution rather than a Gaussian distribution. Under these circumstances the l_1 norm would be more appropriate for describing observation errors. The advantage of using the l_1 norm is that the solution is more robust than when the l_2 norm is used for statistical estimation. Specifically, compared with the l_2 norm a small number of outliers would have less influence on the solution [16,25].

When l_1 norm is used for the deviation from the observations and the prior (sparsity) term a Least Absolute Deviation (*LAD*) solution is obtained:

$$\min_{x \in R^N} \| C_y^{-1/2}(y - \Phi\tilde{x}) \|_1 + \gamma \| \mathbf{W} \cdot \tilde{x} \|_1 \tag{2.10}$$

The LAD formulation can be reduced to a linear programming (LP) problem when the following change of variables is applied (assuming **W** is invertible):

$$u = \gamma \cdot \mathbf{W} \cdot \tilde{x} \quad \text{and}$$

$$v = C_y^{-1/2}(y - \Phi\tilde{x}) \quad \text{or} \quad v = C_y^{-1/2}(y - \gamma^{-1}\Phi\mathbf{W}^{-1}u) \tag{2.11}$$

and by splitting u and v into their nonnegative and nonpositive parts:

$$u = u^+ - u^- \quad \text{and} \quad v = v^+ - v^- \tag{2.12}$$

where $u^+ = \max(u,0)$, $u^- = \max(-u,0)$, $v^+ = \max(v,0)$, $v^- = \max(-v,0)$. The equivalent linear programming formulation can now be written as:

$$\min_{u^+,u^-,v^+,v^-} \mathbf{1}^T u^+ + \mathbf{1}^T u^- + \mathbf{1}^T v^+ + \mathbf{1}^T v^- \tag{2.13}$$

$$\text{subject to } C_y^{-1/2}(y - \gamma^{-1}\Phi\mathbf{W}^{-1}(u^+ - u^-)) = v^+ - v^-$$

$$u^+, u^-, v^+, v^- \geq 0$$

2.4.3 Least Mixed Norm (LMN) Solution ($q = 2,\ p = 1$)

A third alternative formulation is used when the Laplace prior distribution and Gaussian additive observation noise are assumed. In this case, the l_1 and l_2 norms are more appropriate for the prior (sparsity constraint here) and observation deviation, respectively.

$$\min_{x \in R^N} \| C_y^{-1/2}(y - \Phi\tilde{x}) \|_2 + \gamma \| W \cdot \tilde{x} \|_1 \tag{2.14}$$

This formulation is closely related to the total variation regularization that is used in many inverse problem applications [20]. This LMN problem can be written as a quadratic programming (QP) algorithm with the following manipulations:

$$u = \gamma \cdot W \cdot \tilde{x} \ \rightarrow\ (\tilde{x} = \gamma^{-1} \cdot W^{-1} \cdot u) \ \text{ and } \ u = u^+ - u^-$$

$$v = C_y^{-1/2}(y - \Phi\tilde{x}) \ \text{ or } \ v = C_y^{-1/2}(y - \gamma^{-1}\Phi W^{-1}u) \tag{2.15}$$

This reduces the LMN problem to the (QP) formulation below:

$$\min_{u^+, u^-, v} \quad 1^T u^+ + 1^T u^- + v^T v \tag{2.16}$$

$$\text{subject to } \ C_y^{-1/2}(y - \gamma^{-1} \cdot \Phi W^{-1}(u^+ - u^-)) = v$$

$$u^+, u^-, v \geq 0$$

which can be written in standard (QP) form. After eliminating the linear constraint (and variable v) the following (QP) problem is obtained:

$$\min_{x} \ \frac{1}{2}\mathrm{x}^T\mathbf{G}\mathrm{x} + c^T\mathrm{x} \quad \text{subject to} \quad \mathrm{x} \geq 0 \tag{2.17}$$

Where x, G, and c are defined by:

$$\mathrm{x} = \begin{bmatrix} u^+ \\ u^- \end{bmatrix} \ , \ \mathbf{G} = \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix} \ , \ \mathrm{c} = \left[(\mathbf{1}^T - 2\mathrm{r}) \quad (\mathbf{1}^T + 2\mathrm{r})\right] \tag{2.18}$$

$$Q = 2\gamma^{-2}(\mathbf{\Phi}\,\mathbf{W}^{-1})^T C_y^{-1}(\mathbf{\Phi}\,\mathbf{W}^{-1})$$

$$r = \gamma^{-1}\mathrm{y}^{\mathrm{T}} C_y^{-1}\mathbf{\Phi}\,\mathbf{W}^{-1}$$

It is straightforward to verify that the Hessian matrix G is positive-semidefinite and therefore the above objective function is convex (sum of two convex norms). Efficient minimization algorithms are available to solve the above (QP) problem [27]. In the examples of Chapter 3 the solution is obtained using MATLAB (QP) solver *quadprog* [28].

2.5 Incorporation of Dynamic Measurements

The compressed sensing formulations presented in previous section were only solving interpolation problems where dynamic flow equations were not included. While interpolation covers a broad range of inference problems, the goal of this research is to extend the approach to dynamic systems such as those described by equations (1.1)-(1.3).

In order to include the dynamic information into the inference problem, a parameter estimation approach can be formulated that minimizes an objective function comprising

penalty terms for deviation of predicted quantities from observed values, and a term accounting for departure from the prior knowledge about the parameters:

$$\min_{\tilde{x}} J(\tilde{x}) = \| \mathbf{C}_d^{-1/2}(h(\tilde{x}) - d_{obs}) \|_2 + \| \mathbf{C}_y^{-1/2}(\Phi\tilde{x} - y) \|_2 + \| \mathbf{C}_{\tilde{x}}^{-1/2}(\tilde{x} - \bar{x}) \|_2$$

$$\text{subject to } g(\tilde{x}) = 0 \tag{2.19}$$

where, $\tilde{x}$ and $\bar{x}$ are vectors of unknown and prior parameter mean, respectively, with prior covariance specified as $\mathbf{C}_x$; $J(x)$ is the minimization objective function; $h(\tilde{x})$ is the measurement operator that maps parameter space to observation subspace, d_{obs} is the uncertain observation vector with covariance $\mathbf{C}_d$, and $g(\tilde{x})$ represents a discretized version of the multiphase flow equations (1.1)-(1.3), ($g(\tilde{x})$ depends on several other input parameters that are assumed known and dropped here). Since flow equations are derived from mass and momentum conservation principles they are commonly incorporated as *hard* equality constraints. This leads to an adjoint formulation of the problem, that can be solved using variational calculus [29]. The solution of the adjoint model is used to derive the sensitivities of the objective function to unknown parameters (i.e. gradient information). The resulting optimization is a nonlinear non-convex problem that may have several local solutions.

With linear models (h), Gaussian prior distribution, and additive Gaussian error statistics [19,25,30], this objective function can be derived from Baysian probabilistic framework

where the solution yields the peak of the *a posteriori* distribution. When the model is nonlinear this estimator can only approximate the mode of the *a posteriori* distribution.

While a Gaussian assumption for observation error statistics may be justified, prior information may not have a Gaussian distribution. In that case the last term on the right hand side of (2.19) may be a poor choice as the mean $\bar{x}$ may not be a good central estimate of the prior distribution (for instance, when prior is bimodal). In this work, the sparsity constraint is used to include the prior information and improve the ill-posedness of the inverse problem. The proposed parameter estimation approach includes a penalty term in the objective function that accounts for the sparsity of the solution and contains prior information about the unknown parameters. The new objective function can be written as:

$$\min_{\tilde{x}} \; J(\tilde{x}) = \| \mathbf{C}_d^{-1/2}(h(\tilde{x}) - \mathbf{d}_{obs}) \|_2 + \gamma_1 \| \mathbf{C}_y^{-1/2}(\Phi\tilde{x} - y) \|_2 + \gamma_2 \| \mathbf{W} \cdot \tilde{x} \|_1$$

$$\text{subject to} \quad g(\tilde{x}) = 0 \tag{2.20}$$

The two parameters γ_1 and γ_2 control the relative significance that is given to each term in the objective function. The first two departure terms are penalized using l_2 norm while the sparsity constraint is enforced through a l_1 norm (this choice is based on experiments with different formulations presented in previous section and is discussed in the next chapter).

Several first order search algorithms can be used to solve the resulting optimization problem [31,32]. However, an efficient approach that has been successfully used to solve similar optimization problems is the Limited Memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) quasi-Newton (or its limited memory version for large problems) [33]. The adjoint solution resulting from the mass conservation constraint can be used to compute the gradient information efficiently [8]. The implementation of the adjoint model to derive the gradient information is available in the commercial reservoir simulator used in this study [8]. It is worthwhile to note that the gradient information is obtained with respect to the original permeability field, which is then converted to the transformed permeability field (DCT domain) through a simple differentiation chain rule.

The approach proposed here is expected to offer an overall improvement in computational cost, ill-posedness, and accuracy of the original formulation of the inverse problem (2.19). This inverse modeling framework is implemented in Chapter 4 using a waterflooding experiment with a synthetic 2D reservoir model.

Chapter 3

Results and Discussion

In this chapter, application of the three approaches introduced in Chapter 2 to an interpolation problem with a channelized permeability field is presented. The purpose of this chapter is to investigate the form of sparsity constraint that should be used for inversion. The experimental setup is described in Section 3.1. The problem formulations that resulted in *LAD*, *LLS* and *LMN* solutions are discussed in Section 3.2. Section 3.3 discusses a more realistic setting, in which fewer observations are used and observation locations are fixed. Finally, the chapter is closed with a brief discussion on the overall results and concluding remarks.

3.1 Experimental Setup

This chapter investigates the application of the general problem formulations in Chapter 2 with an example of channelized reservoir permeability. The example is similar to that in Section 2.2. Two sets of experiments are designed each with a different purpose.

In the first set of experiments, suitability of the three problem formulations in Chapter 2 is studied. In each experiment, four different test runs are considered. Sensitivity of each of these formulations to two conditions is tested: 1) training DCT basis with the prior library; and 2) sparsity of the original permeability. Varying the above parameters results in four test runs:

T1: No prior basis training and the original permeability field is perfectly sparse

T2: With prior basis training and the original permeability is perfectly sparse

T3: No prior basis training and the original permeability field is not sparse

T4: With prior basis training and the original permeability is not sparse

In these set of runs the observation points are randomly selected. The observation locations and their values are identical in each of these runs. The sparsity level, dimension of search subspace, and the number of observations is fixed at $S=15$, $N=78$, and $K=30$. It is important to note that the sparsity level $S = 15$ is only relevant when the original permeability field has been made sparse through truncation in the DCT domain (*T1* and *T2* cases above). When trained, the prior weighting matrix $\mathbf{W}$ is obtained from

the same permeability library and is fixed in all experiments. Otherwise, this matrix is set equal to identity. The value of parameter γ is varied to arrive at a balance in the contributions from observations and the sparse prior in the objective function that resulted in the best achievable estimate. Based on these tests, a formulation that gives a robust and accurate performance is selected for application in a more realistic setting described below.

A second set of experiments is designed to investigate the performance of the selected formulation in previous tests under more realistic assumptions. In these experiments, the observation locations are fixed at the two (east-west) ends of the reservoir. This setting will also be used for dynamic data integration in a waterflooding experiments in Chapter 4. The number of observations in these experiments is changed to $K = 90$ while the search dimension and sparsity level are the same as above ($N=78$ and $S=15$). In these experiments the sensitivity to $\gamma$is also studied. Three test runs are performed with small, intermediate, and large γvalues corresponding to non-sparse, moderately sparse, and very sparse estimates, respectively. These experiments are conducted with and without including prior basis training.

The results of the experiments described above are summarized in the next section.

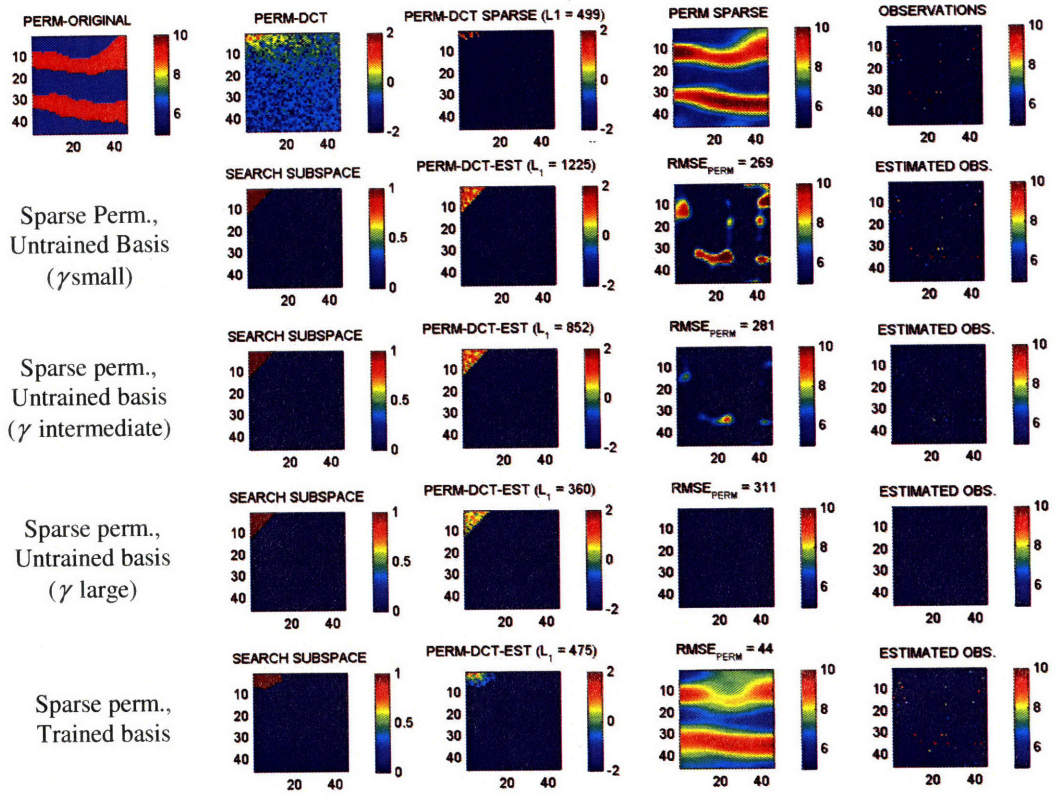
3.2 Interpolation Using Spatially Random Observations

3.2.1 Linear Least Squares (*LLS*) Solution

The solutions to the *LLS* formulation for the experimental setup with random observations are shown in Figure 3.1. In Figure 3.1a when no basis training is used (second, third, and fourth rows) it was not possible to find a reasonable estimate for the sparse permeability field. Rows two, three, and four in Figure 3.1a show the estimation results by specifying a small, intermediate, and large weight for sparsity term in the objective function, respectively (through adjusting γ). Similar results are achieved for the case when the original permeability field is not perfectly sparse in Figure 3.1b. These results may be expected due to the fact that the l_1 norm provides a better measure of sparsity (it is closer approximation to the original l_0 norm that defines the exact number of sparse coefficients).

The last rows in Figures 3.1a and 3.1b illustrate the *LLS* estimates when prior basis training is used. As seen in these figures the *LLS* estimates perform well only when an accurate prior is used to weigh the appropriate basis vectors. In the absence of a good prior, the l_2 norm constraint that is used for sparsity does not yield the channelized structure observed in the original permeability. Closer examination of the results reveals that the estimated DCT coefficients are not sparse confirming that the l_2 norm is not a good choice to preserve sparsity.

a) *LLS* reconstruction results for an originally sparse permeability



b) *LLS* reconstruction results for an originally non-sparse permeability

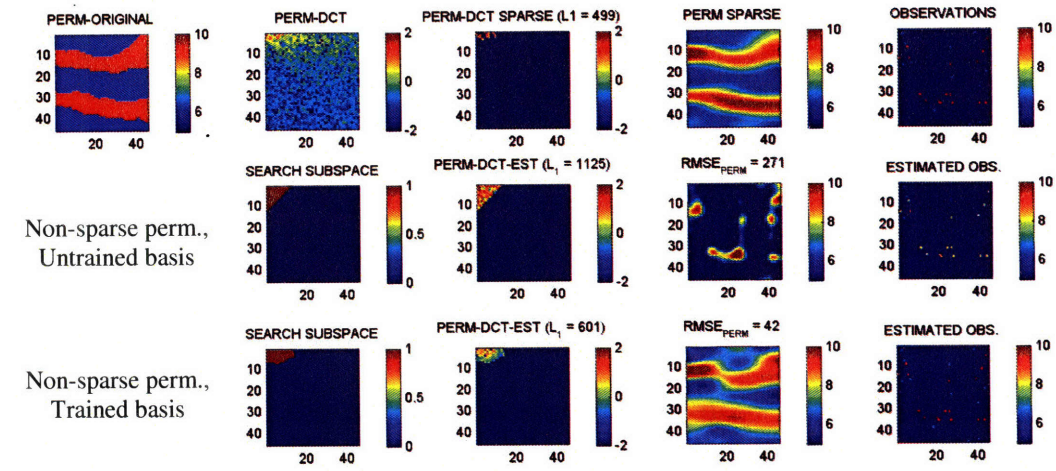


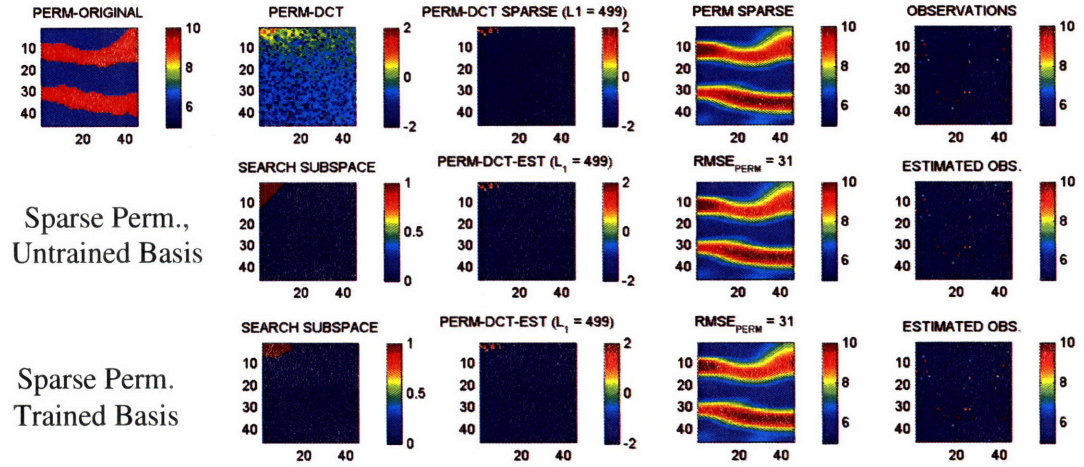
Figure 3.1 *Linear Least Squares* (*LLS*) reconstruction results for originally sparse (a) and non-sparse (b) permeability fields: the first row in (a) and (b) show (left to right) the original permeability field, its DCT transformations, the truncated DCT with $S = 15$ non-zero coefficients, reconstructed sparse permeability field, observations used in the estimation (observations in (a) and (b) are generated from sparse and non-sparse permeability fields, respectively); the second to last rows show (left to right) the search subspace, estimated DCT coefficients, estimated permeability field, and estimated observations; rows two to five in (a) show the estimation results under T1 experimental conditions with increasing sparsity level (top to bottom).

These results suggest that the l_2 norm fails to identify the sparsity of the original signal and can only provide a reasonable estimate when accurate prior information is supplied. In general, l_2 norm is very sensitive to large deviations (outliers) and tends to ignore smaller terms, which is not desirable for preserving sparsity. This can have negative implications in realistic settings, in which the prior information may be inaccurate.

3.2.2 Least Absolute Deviation (*LAD*) Solution

The solutions to the *LAD* formulation for similar experimental setup as in Figure 3.1 are shown in Figure 3.2. Figure 3.2a shows the case for which the original permeability field is made sparse through the truncation shown in the first row. Two sets of estimates are shown in this figure. The second row in Figure 3.2a shows estimates with no basis training and the third row contains the results with prior basis training. It is seen from these results that the l_1 norm is very effective in perfect reconstruction of the original sparse signals. It is important to note that these results are a function of the sparsity S, search dimension N, and number of observation K. The results of these experiments suggest that for the chosen values of S, N, and K the untrained bases are sufficient for constructing the original sparse signal. This is not the case when a smaller K is used and the trained basis is expected to give better reconstruction if a reasonable training library is used.

a) *LAD* reconstruction results for an originally sparse permeability



b) *LAD* reconstruction results for an originally non-sparse permeability

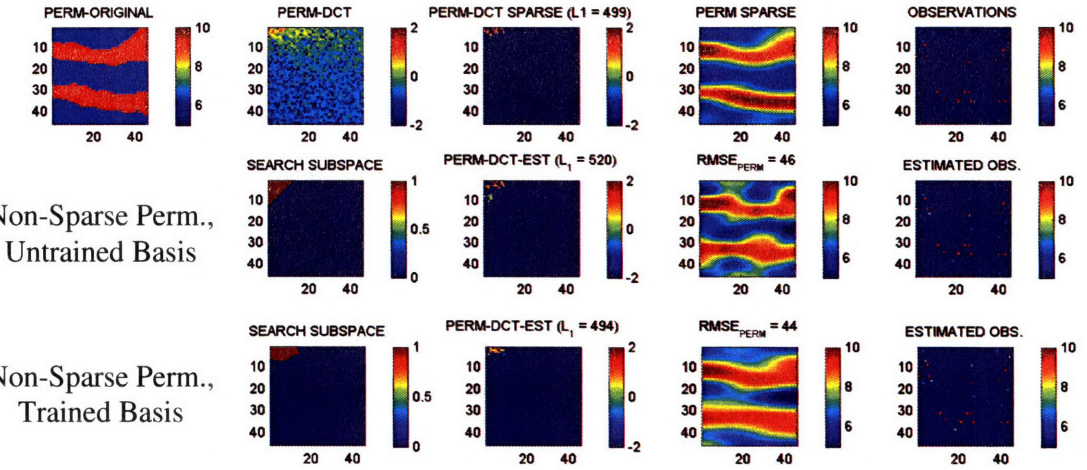


Figure 3.2 *Least Absolute Deviations (LAD)* reconstruction results for originally sparse (a) and non-sparse (b) permeability fields: the first row in (a) and (b) show (left to right) the original permeability field, its DCT transformations, the truncated DCT with $S = 15$ non-zero coefficients, reconstructed sparse permeability field, observations used in the estimation (observations in (a) and (b) are generated from sparse and non-sparse permeability fields, respectively); the second and third rows show (left to right) the search subspace, estimated DCT coefficients, estimated permeability field, and estimated observations.

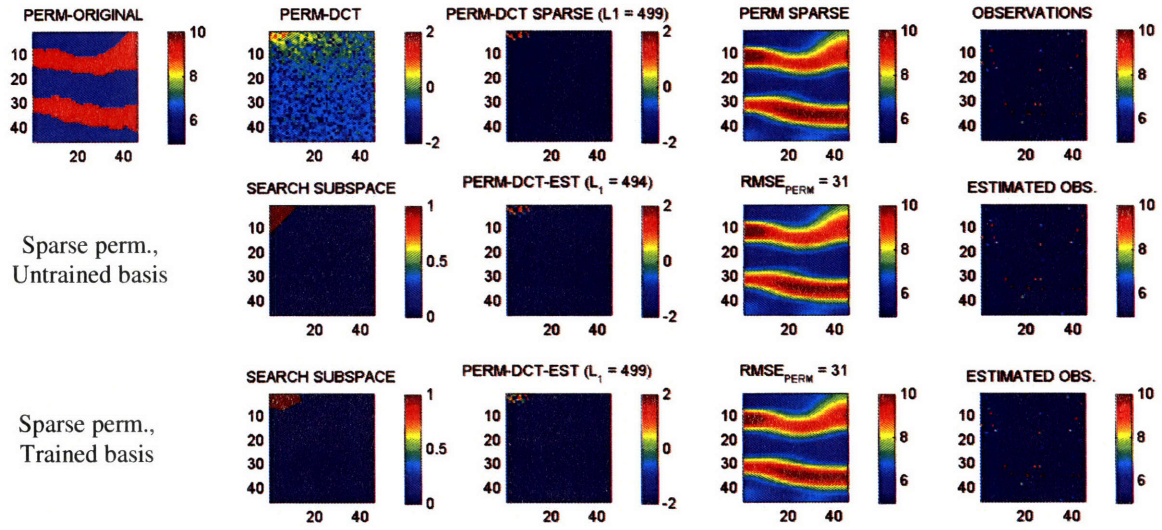
Figure 3.2b illustrates the same experiment as in 3.2a, except that the original signal has not been made sparse. One way of interpreting this effect is assuming that the original signal is sparse as in Figure 3.2a but the observations are noisy (observations come from the original non-sparse permeability field). As seen in Figure 3.2b, the results are quite sensitive to observation noise, however, the trend in the permeability field is captured.

3.2.3 Least Mixed Norm (*LMN*) Solution

The results in Sections 3.2.1 and 3.2.2 imply that for a robust reconstruction with and without prior training of the basis an l_1 norm representation of the sparsity constraint is more appropriate. This is tested with the implementation of the *LMN* formulation for similar examples. Figure 3.3 summarizes the results for *LMN* reconstruction in a similar manner to previous sections. As seen in second rows of Figures 3.3a and 3.3b, unlike the *LLS* reconstruction results, the *LMN* reconstruction is quite accurate when prior basis training is not used. This further confirms the conclusion of the previous section, that the sparsity constraint is important in the reconstruction and better preserved using an l_1 norm.

Figure 3.3b shows the *LMN* reconstruction results for the case in which the original signal is not sparse. The results appear to have similar performance to those of the *LAD* reconstruction solutions. While the accuracy of the two methods is similar, the *LMN* formulation requires solution of a quadratic programming (*QP*) problem, which can be computationally more demanding than solving a linear programming (*LP*) problem in the *LAD* formulation. While this distinction is very important to make for the

a) *LMN* reconstruction results for an originally sparse permeability



b) *LMN* reconstruction results for an originally non-sparse permeability

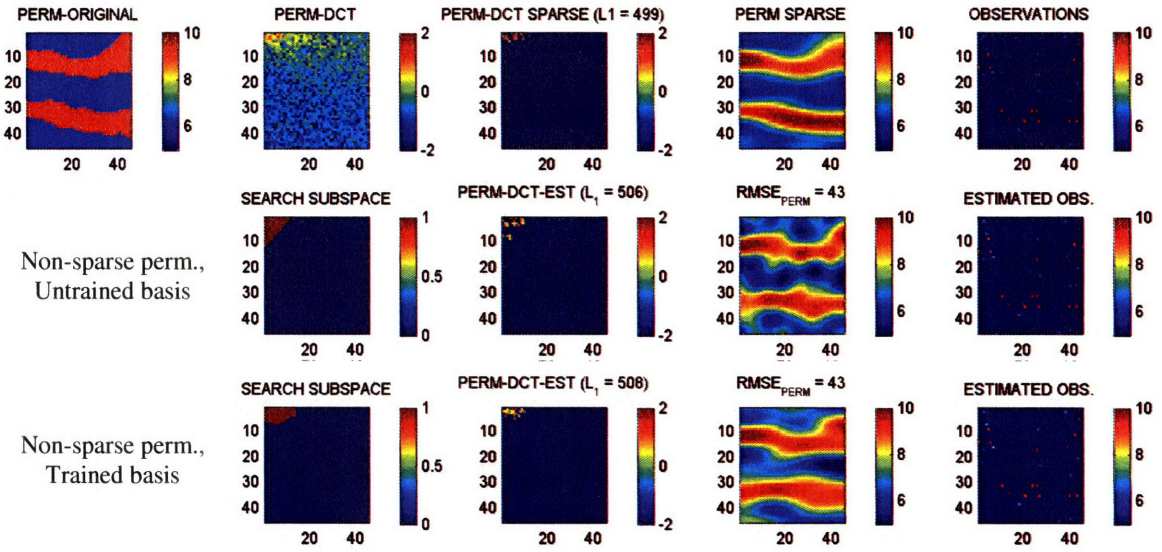


Figure 3.3 *Least Minimum Norm* (*LMN*) reconstruction results for originally sparse (a) and non-sparse (b) permeability fields: the first row in (a) and (b) show (left to right) the original permeability field, its DCT transformations, the truncated DCT with $S = 15$ non-zero coefficients, reconstructed sparse permeability field, observations used in the estimation (observations in (a) and (b) are generated from sparse and non-sparse permeability fields, respectively); the second to last rows show (left to right) the search subspace, estimated DCT coefficients, estimated permeability field, and estimated observations.

interpolation example, in the dynamic data integration problem with a non-linear flow model the *LP* and *QP* formulations are not possible and an adjoint-based formulation is used to minimize the corresponding objective function.

Given the results in of three formulations in this section, the *LMN* formulation is adopted for further analysis and development with spatially fixed observation points in the next section.

3.3 Spatially Fixed (Non-Random) Observations

The examples in Section 3.2 used observations that were randomly located in space. This, however, is not likely to happen in operational settings. The point observations in real reservoirs come from the drilled production, injection, and monitoring wells. The prohibitive cost associated with drilling activities result in minimal number of wells that are drilled in specified locations. This raises two issues that need to be addressed.

First, the non-random nature of well locations may conflict with the assumptions and results of Section 3.2. Depending on the abundance of data in the field (the second issue to be addressed next) this may or may not be a problem. In general, limited localized data tend to degrade the performance of the compressed sensing formulation. The randomness requirement of the observations would allow for wider signal support coverage in its non-sparse domain, which provides more information content and helps its reconstruction in the domain in which it is sparse. This can potentially become an issue in field settings

where limited observation points are available. However, the production wells are usually spread in the field and provide a good coverage.

The second issue to address is the quantity of the observations. This is likely to be a limiting factor in practice if interpolation problems with limited well locations are considered. Fortunately enough, there are other sources of measurements that can complement static point observations. Examples of these measurements are 3D and 4D seismic data, well logs, and dynamic production data. Availability of each of these data sources can further constrain the reconstruction problem and compensate for the observational requirement of the reconstruction formulations in Chapter 2. In this section, reconstruction of the permeability field with fixed observation locations is considered. Integration of additional dynamic production measurements is considered in Chapter 4.

3.3.1 *LMN* Solution with Spatially Fixed Observations and Untrained Basis

In this section, the reconstruction results with spatially fixed observation points and without including the prior information is presented. The observations are located at the east-west ends of the reservoir to resemble a horizontal waterflooding scenario in a smart oilfield setting (Figure 1.1). In this setup, the observation points are localized, which is expected to adversely affect the reconstruction.

Figure 3.4 shows the reconstruction results when the original permeability is made sparse ($S=15$) and with increasing confidence in sparsity, forced by adjusting the value of γ(the rows from top to bottom). It is clear from these results that the observations are not

sufficient to allow for reconstruction of the original permeability using sparsity constraint. It is important to note that the number of observations here ($K=90$) is three times the number of observations in the examples of Section 3.2. However, the observations are not randomly located and fail to provide global information. The increasing degree of sparsity constraint as seen in rows three and four of Figure 3.4 manifest itself in the extension of the high permeability observation regions inward, but a sparse solution appears to exist that provides a smaller objective function than the true permeability. Similar results are obtained when the original permeability is non-sparse (not shown).

***LMN* reconstruction results for spatially fixed observation points (untrained basis)**

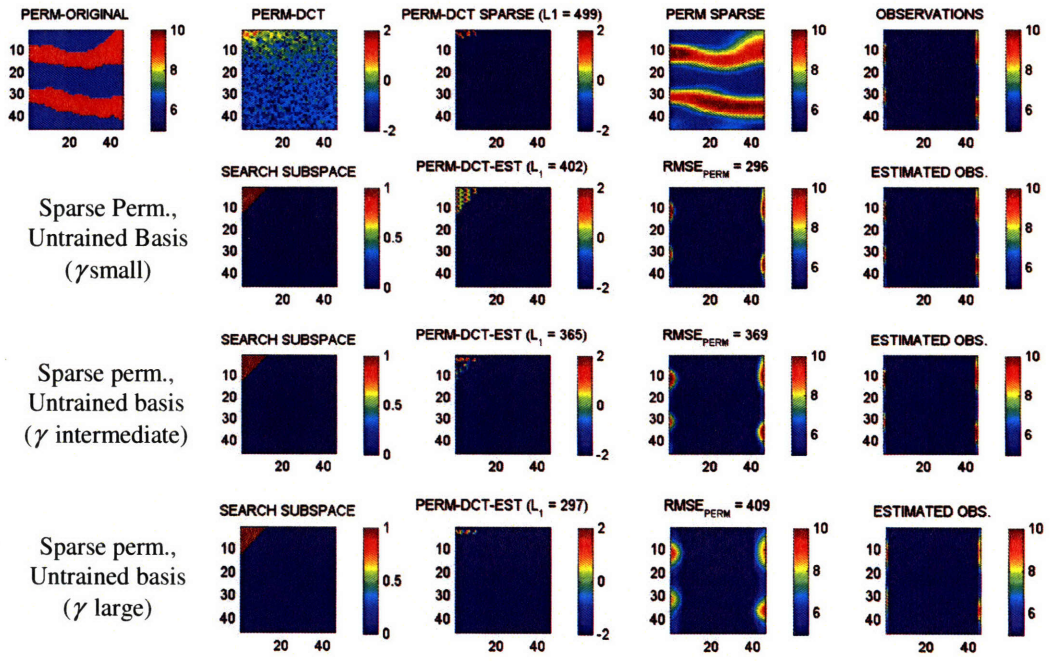


Figure 3.4 *Least Mixed Norm* (*LMN*) reconstruction results for originally sparse permeability field using spatially fixed observation points and untrained basis: rows two to four show the estimation results with increasing sparsity level (top to bottom).

3.3.2 *LMN* Solution with Spatially Fixed Observations and Trained Basis

The *LMN* reconstruction results of previous section can be improved by training the basis vectors. This assumes a good prior library is available and is expected to make the solution sensitive to prior specification (less robust).

Figures 3.5 and 3.6 summarize the reconstruction results for the cases in which the original permeabilities are sparse and non-sparse, respectively. As expected, incorporating the prior information in the basis guides the reconstruction algorithm in identifying the two horizontal channels.

***LMN* reconstruction results for spatially fixed observation points (prior-trained basis)**

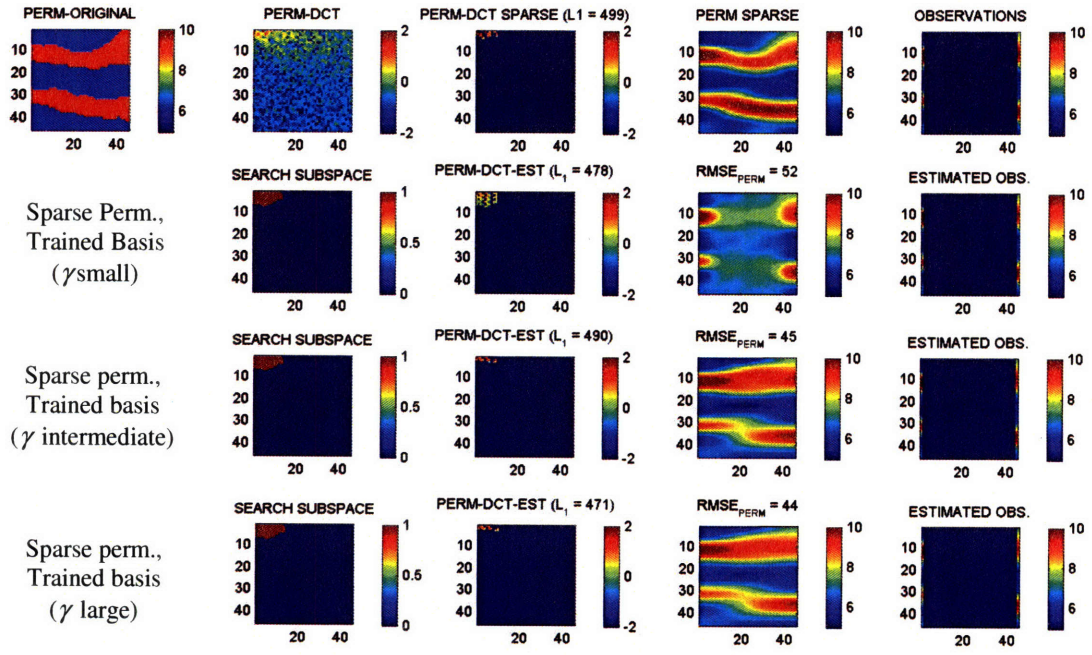


Figure 3.5 *Least Mixed Norm* (*LMN*) reconstruction results for originally sparse permeability field using spatially fixed observation points and prior-trained basis: rows two to four show the estimation results with increasing sparsity level (top to bottom).

Although the prior-trained reconstruction identifies the presence of channels in Figures 3.5 and 3.6, the exact shape of the channels is not captured well. This implies that even when correct prior information is included in the reconstruction, the sparsity constraint may not result in accurate identification of the existing features if sufficient observations are not available to constrain the algorithm. The example in Figure 3.6 shows similar results for the case in which the original signal is not sparse (or the observations are noisy). The results suggest that the solution still identifies the two channels under prior-trained basis and sparsity constraint. Again in this case, however, the channels are not accurately retrieved due to insufficient information content in the observations. This calls for integration of additional observations to constrain the reconstruction algorithm, which is considered in Chapter 4.

***LMN* reconstruction results for spatially fixed observation points (prior-trained basis)**

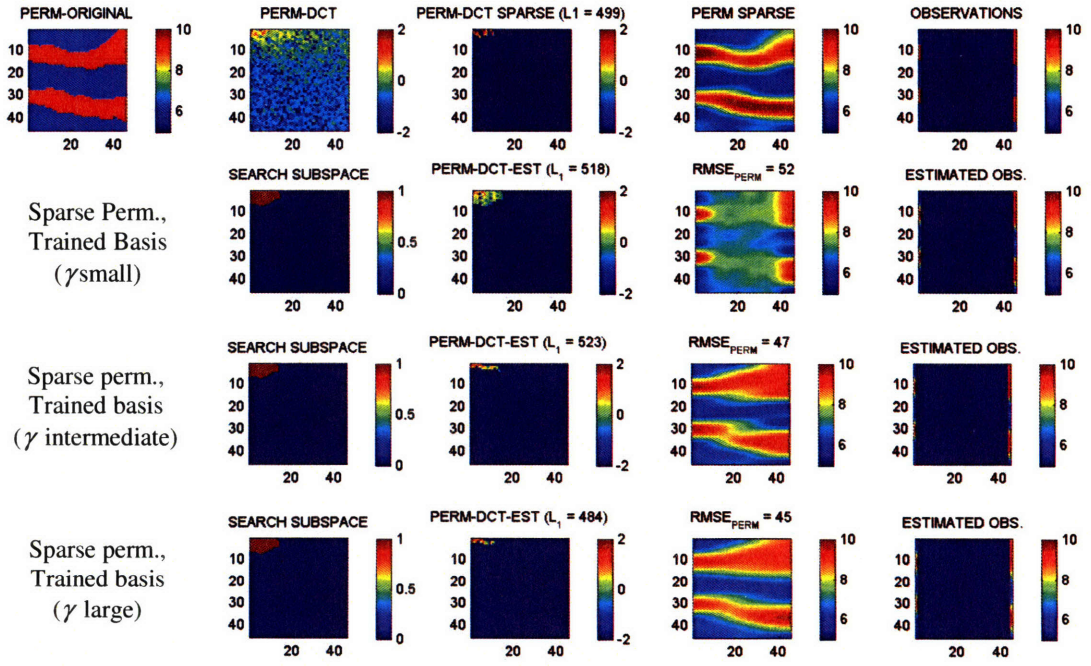


Figure 3.6 *Least Mixed Norm* (*LMN*) reconstruction results for originally *non-sparse* permeability field using *spatially fixed observation* points and *untrained basis*: rows two to four show the estimation results with increasing sparsity level (top to bottom).

3.4 Sensitivity to Observation Errors

The results presented in previous sections assumed perfect observations, which is often not the case in practice. In this section sensitivity of the estimates to different levels of measurement errors is studied. The ill-posed nature of the problem necessitates such sensitivity studies to assess the robustness of the reconstruction formulation to the noise level in the measurements. Figure 3.7 shows the estimation results for *LMN* estimate in a similar experiment to that shown in Figure 3.6. In this example additive Gaussian pseudo-random noise with zero mean and standard deviations equivalent to 5%, 10%, 20%, and 50% of the observation mean are considered in the reconstruction.

From Figure 3.7 reconstruction results appear to be quite robust for noise levels less than 20% while degradation in the quality is observed when higher noise levels are considered. At noise levels of 50% or more the reconstruction loses the location of the present channels in the field. However, for this application, since point measurements in the field have very small errors associated with them, higher level of measurement errors does not seem to be a major concern. This may become an important issue when remotely sensed seismic observations with higher levels of uncertainty are integrated.

Sensitivity of the reconstruction to level of noise in observations

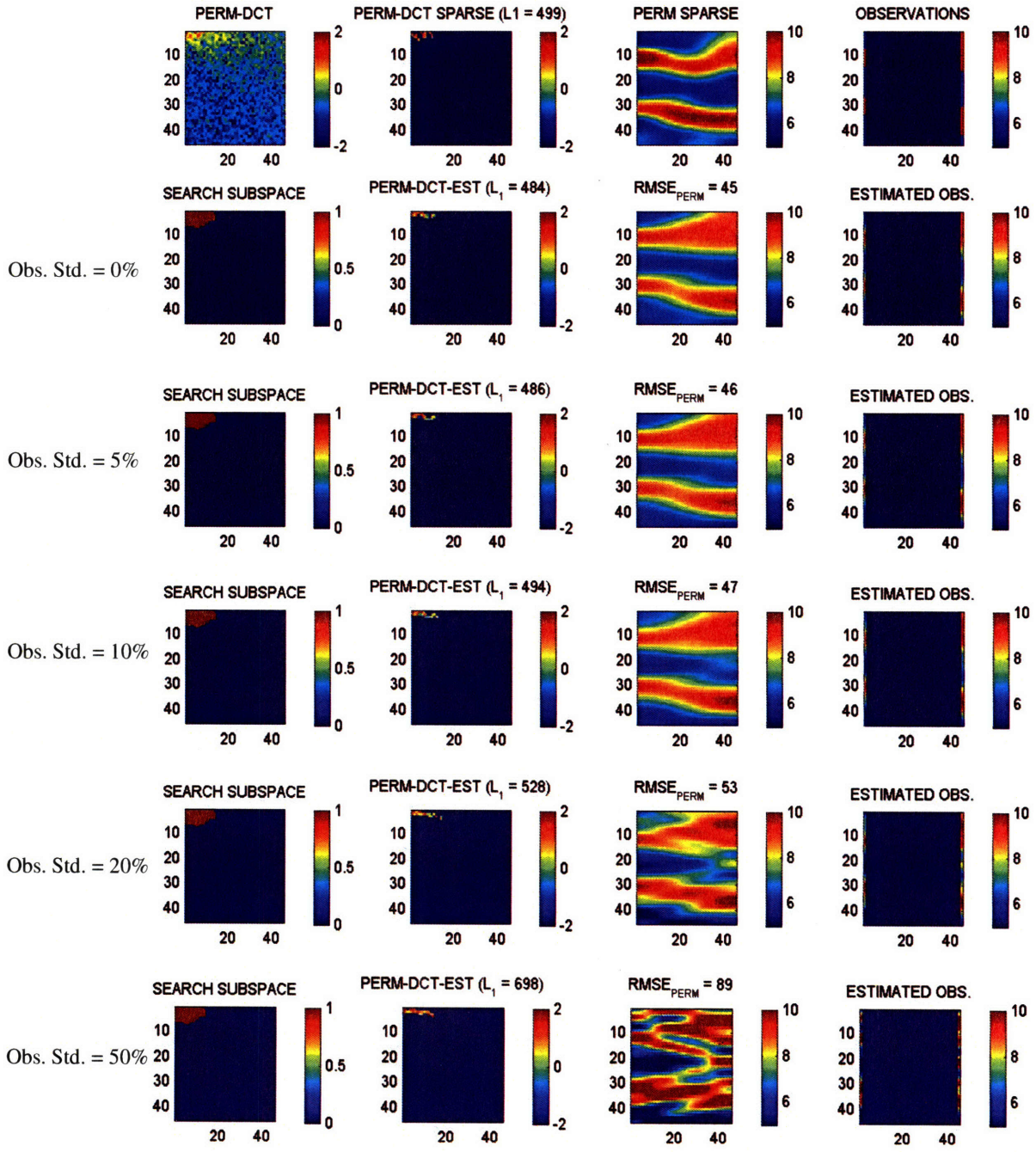


Figure 3.7 Sensitivity of the *LMN* reconstruction results to errors in the measurements: effect of different levels of measurement errors (0%, 5%, 10%, 20%, and 50%) on the quality on the estimates (rows two to six, respectively) is depicted. The degradation in the quality of the estimates is most pronounced for noise levels above 20% of the mean values.

3.5 Conclusions

The results presented in this chapter leads to the conclusion that the l_1 norm can represent the sparsity constraint in the reconstruction of sparse (or approximately sparse) signals while the l_2 norm fails to do so without additional assumptions. The l_2 norm could provide reasonable reconstruction results only when accurate prior information was included. For the sparsity constraint formulation to perform well, it is also essential to constrain the algorithm with observations that provide sufficient information content (signal support coverage). The examples suggested that a large number of localized observations may constrain the reconstruction algorithm less than fewer observations that are distributed through the domain and provide a good global coverage. This may or may not be the case in realistic reservoirs depending on the field type and production strategies; however, it is an important consideration in applying the proposed approach.

While in realistic reservoirs point observations can be limited due to prohibitive data acquisition costs, other sources of information may be present to constrain the reconstruction algorithm. One of the most frequently used observation types is the dynamic production information. These observations are obtained during the development phase of the reservoir and indicate reservoir's response to specified injection/production strategies at well locations. This is investigated further in Chapter 4.

Chapter 4

Inversion Using Dynamic Observations

This chapter discusses the application of the methodology proposed in previous chapters to an inverse problem with time-variant state-space model. The model describes displacement of two immiscible fluids (oil and water) in subsurface media by solving the system of equations in (1.1)-(1.3). The chapter starts with a description of the experimental setup and the water-flooding example that will be used in the inversion study. The inversion formulation and an overview of the solution method is discussed next, followed by the presentation of several examples that are used to illustrate the importance of prior training versus sparsity constraint.

4.1 Experimental Setup

A water flooding example that is also studied in [4,34] is considered in this section. In this example a 450 m × 450 m × 10 m synthetic reservoir is discretized into a two-dimensional 45 × 45 × 1 uniform grid block system, as shown in Figure 1.1. The simulations are performed with the commercially available ECLIPSE [8] reservoir simulator, which is set up for two phase (oil and water) black oil flow. Observations after 90 days of simulations are integrated as dynamic production data. Horizontal wells with 45 ports are used to inject water uniformly into the left side of the reservoir and to produce oil and water from the right side end. The injection wells are operated with specified flow rates while the production wells are operated with specified pressures (2950 psi).

In this study the only source of simulator uncertainty is the permeability, which is treated as a random field. Initial and boundary conditions are assumed to be known perfectly and dynamic model errors are assumed to be negligible. The initial reservoir pressure and water saturation are 3000 psi and 0.10, respectively, throughout the reservoir. Two types of measurements are assumed to be available: 1) pressure observations at each of the 45 ports in the injection wells and 2) oil and water flow rate measurements at each of the 45 ports in the production wells. In each experiment the observations of injection well pressures and production well flow rates are generated by running the simulator from a specified "true" permeability field. After integrating the observations on day 90, the permeability estimates are used to perform flow simulations for 360 days to predict future saturation profiles and compare them with the true saturations.

4.2 Problem Formulation with Dynamic Observations

The compressed sensing framework and examples presented in previous chapters were only for interpolation problems and dynamic observations were not included. In addition, it was clear from examples in Section 3.3.2 that when the number of observations is reduced and observation locations are fixed, the theoretical requirements of the compressed sensing algorithm may not be met and the reconstruction can become under-constrained. However, in dynamic systems time-variant information may be available to help further constrain the reconstruction problem.

Dynamic observations usually provide additional information about the response of the system to specified input forcing. While the response may be observed in a particular location, it is often lumped and contains regional or even global information about existing features in the system. Therefore, dynamic observations can often provide information about a system that goes beyond the measurement location and can be used to make inferences about the unknown global attributes of the system.

In dynamical systems, a parameter estimation approach can be formulated that include sparsity constraint, prior information, and observation of static and dynamic attributes of the system as discussed in Section 2.5. The resulting objective function is given below:

$$\min_{\tilde{\tilde{x}}} J(\tilde{x}) = \| \mathbf{C}_d^{-1/2}(h(\tilde{x}) - \mathbf{d}_{obs}) \|_2 + \gamma_1 \| \mathbf{C}_y^{-1/2}(\Phi\tilde{x} - y) \|_2 + \gamma_2 \| \mathbf{W} \cdot \tilde{x} \|_1$$

$$\text{subject to } g(\tilde{x}) = 0 \tag{4.1}$$

The nonlinear constraint *g(.)* represents mass and momentum conservation principles that are used in deriving equations (1.1)-(1.3). Adjoint-based optimization methods can be used to derive the normal equations of the augmented objective function and compute, quite efficiently, its gradient with respect to unknown parameters in spatial domain. Gradient-based search algorithms that can be used to minimize objective functions of this type have been reported in the literature [31,32]. A particularly suitable search method for the problem of interest in this section is the quasi-Newton LBFGS algorithm [31,32].

The adjoint solution is implemented in the commercial reservoir simulator used in this research and the LBFGS implementation in MATLAB's *fmincon* [28] has been used to find the solutions. It is important to note that the output gradients from ECLIPSE are computed for well variables (i.e *h(x)*) with respect to the original parameters in the spatial domain. Therefore, the differentiation chain rule has to be used to obtain the Jacobian of the objective function with respect to the retained DCT coefficients as follows:

$$J(\tilde{x}) = (h(\tilde{x}) - \mathbf{d}_{obs})^T \mathbf{C}_d^{-1} (h(\tilde{x}) - \mathbf{d}_{obs}) + \gamma_1 (\Phi\tilde{x} - y)^T \mathbf{C}_y^{-1} (\Phi\tilde{x} - y) + \gamma_2 \| \mathbf{W} \cdot \tilde{x} \|_1 \tag{4.2}$$

$$\mathrm{d}J(\tilde{x}) = \frac{\partial J_1}{\partial h(\tilde{x})} \mathrm{d}h(\tilde{x}) + \gamma_1 \frac{\partial J_2}{\partial \tilde{x}} \mathrm{d}\tilde{x} + \gamma_2 \frac{\partial J_3}{\partial \tilde{x}} \mathrm{d}\tilde{x}$$

$$\frac{\mathrm{d}J(\tilde{x})}{d\tilde{x}} = \frac{\partial J_1}{\partial h(\tilde{x})} \frac{\mathrm{d}h(\tilde{x})}{d\tilde{x}} + \gamma_1 \frac{\partial J_2}{\partial \tilde{x}} + \gamma_2 \frac{\partial J_3}{\partial \tilde{x}} \tag{4.3}$$

Where J_1, J_2, and J_3 refer to the first, second, and third term of the right hand side of the objective function (4.2). In addition, $\frac{\mathrm{d}h(\tilde{x})}{\mathrm{d}\tilde{x}}$ can be computed through the differentiation chain rule and using $\frac{\mathrm{d}h(y)}{\mathrm{d}y}$, which is the gradient information calculated by the simulator:

$$\Phi\tilde{x} = y \Rightarrow \Phi \mathrm{d}\tilde{x} = \mathrm{d}y \Rightarrow \frac{\mathrm{d}h(\tilde{x})}{\mathrm{d}\tilde{x}} = \frac{\mathrm{d}h(\tilde{x})}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}\tilde{x}} = \frac{\mathrm{d}h(\tilde{x})}{\mathrm{d}y}\Phi \tag{4.4}$$

The final result can be written as:

$$\frac{\mathrm{d}J(\tilde{x})}{d\tilde{x}} = 2\mathbf{C}_d^{-1}(h(\tilde{x}) - \mathbf{d}_{obs})\frac{\mathrm{d}h(\tilde{x})}{\mathrm{d}y}\Phi + 2\gamma_1\Phi^T\mathbf{C}_y^{-1}(\Phi\tilde{x} - y) + \gamma_2\mathbf{W}\cdot\frac{\tilde{x}}{\| \tilde{x} \|_1}$$

s (4.5)

All the terms in this Jacobian equation are known and the LBFGS search method can be used to reduce the value of this objective function to its (local) minimum.

A few remarks are in order before ending this section: first, in doing the parameter estimation, the logarithm of the permeability field is usually used instead of the original permeability values. This leads to a slight modification (change of variable) in equations (4.1)-(4.2), which is not shown here. Second, since the LBFGS is a gradient-based search method and the objective function is nonlinear and non-convex, it is likely for the search algorithm to be trapped in a local minimum. Therefore, unlike the LP and QP formulation of previous chapters, a global minimum solution is not guaranteed.

4.3 Estimation/Reconstruction Results

In this section the results for integration of dynamic measurements in the reconstruction with and without sparsity constraint are presented. The cases with trained and untrained bases are also considered separately. First, a solution is presented without including any prior information and where no assumption has been made about the sparsity of the parameter. This is followed by examples in which different levels of prior information are incorporated as discussed briefly here.

First, the prior information can be included through providing a selected set of more relevant basis vectors using the available image library or training images. The new basis vectors contain information about the directionality of the existing features, which can make the solution sensitive to the accuracy of the specified prior. This approach does not include information about sparsity and is used as a reference case to evaluate the performance of the solution with sparsity constraint and trained basis. To address the issue of sensitivity to prior specification, a second method can be considered in which the prior knowledge enters in the form of sparsity constraints without including any information about the directionality of the features present in the library. This approach is particularly useful for situations in which the prior knowledge is either unavailable or inaccurate. However, when dependable prior information is available, the search space can be trained and a more appropriate subset of basis images can be selected, in addition to the sparsity constraint. The results for all of these implementations are presented and discussed in this section.

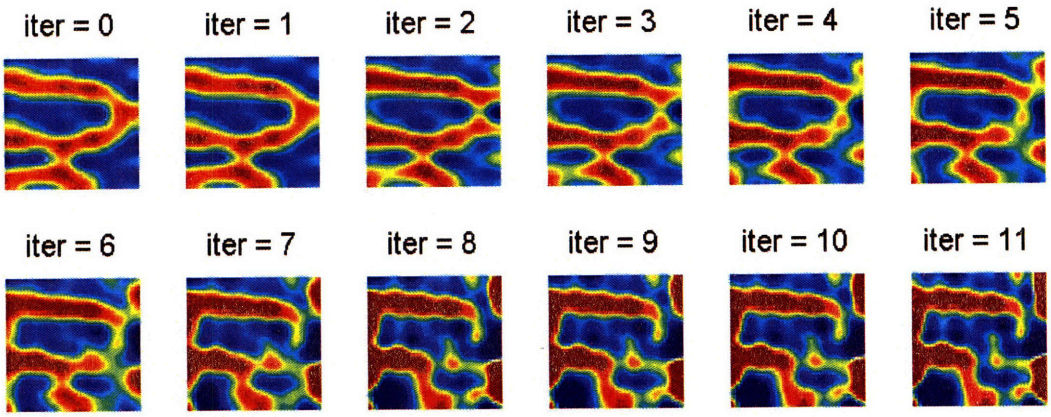
4.3.1 Solution with Untrained Basis and No Sparsity Constraint

The solution discussed in this section is related to the case in which no basis training is performed and the sparsity constraint is not included in the objective function. That is, γ_2 is set equal to zero in equation (4.1) and the prior information has not been used to train DCT basis vectors. Therefore, a solution is sought without providing any directionality preference or weighting for DCT basis vectors. Figure 4.1 shows the reconstruction plots and the reduction in objective function after each iteration in the minimization. The maximum number of minimization iterations was set to 20 as in most cases after 10 iterations no major improvement was observed in the objective function and the estimated parameters. In general, it was observed that the algorithm took some time to eliminate incorrect high permeability regions in the initial field at locations away from the observation areas. Even in the end, small signatures of these regions can be seen in the estimates.

It is clear from Figure 4.1 that the reconstruction algorithm can not identify the connectivity of the channels even though accurate observations from the two ends of the reservoir are available and included. Without including the sparsity constraint and prior information, the reconstruction algorithm appears to have failed to capture the shape of the channels and overestimates (darker red) the values of high permeability areas while it underestimates (darker blue) low permeability regions. These over-/under- estimations can compensate for each other in terms of fluid flow velocity and may produce a net (average) travel time that is consistent with the observations. In the next section, an identical experiment is conducted in which sparsity constraint is added to the objective

function. However, no prior assumption is made about the directionality of the channels and the weighting matrix (**W**) is identity.

(a) Permeability estimates after each minimization iteration



(b) Reduction in objective function after each minimization iteration.

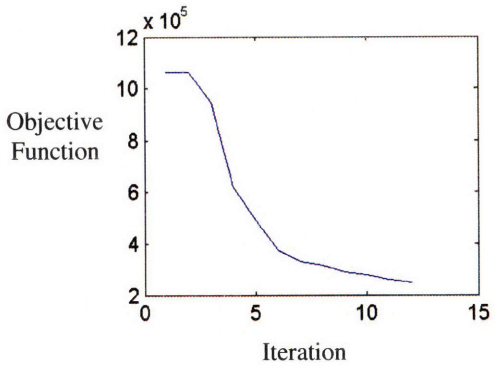


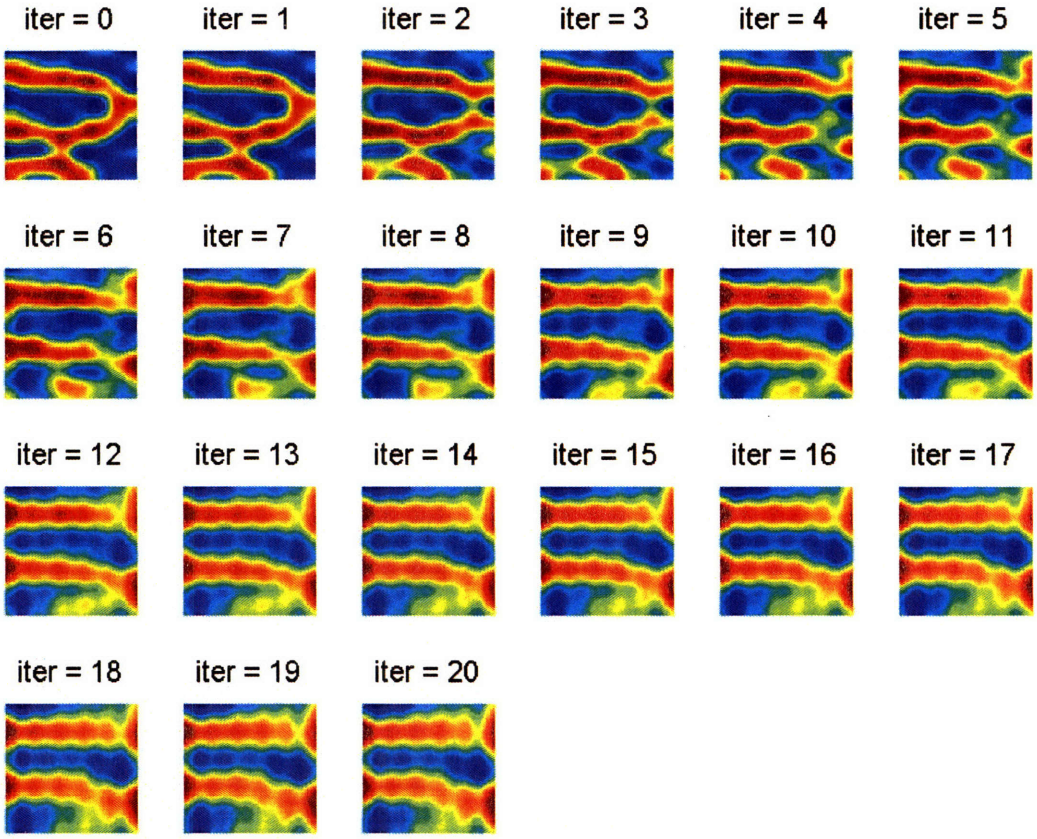
Figure 4.1 Reconstruction results for inversion with static and dynamic measurements without using prior information and sparsity constraint: (a) log-perm after each minimization iteration; (b) reduction in the objective function at each iteration.

4.3.2 Solution with Untrained Basis and Sparsity Constraint

Figure 4.2 shows the solution of the same inverse problem as in Section 4.3.1 when sparsity constraint is added to the objective function. The other terms in the objective function remain identical to the previous example. As seen in Figure 4.2 the sparsity constraint seems to be able to connect the two end of the reservoir where measurements are taken. It is important to note that the algorithm did not use any information about presence of channelized structures and their directionality in the field, nevertheless it was able to approximately detect the shape of the existing channels.

A comparison between Figures 4.2 and 3.4 (similar experiment without dynamic measurements) reveals that the production data have played a major role in capturing the existing channel structure. It is noteworthy that even though these observations were collected at exactly the same locations as static measurements, they provide lumped information about values of the permeability away from the observation points. This result is promising as it indicates the usefulness of the approach in problems where static point observations are limited in space and dynamic measurements may be available to provide additional information about unobserved locations.

(a) Permeability estimates after each minimization iteration



(b) Reduction in objective function after each minimization iteration.

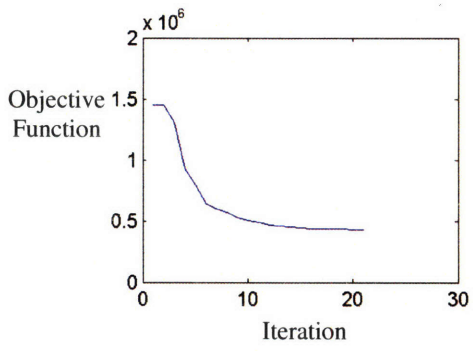


Figure 4.2 Reconstruction results for inversion with static and dynamic measurements without using prior information and when sparsity constraint in included: (a) log-perm after each minimization iteration; (b) reduction in the objective function at each iteration.

4.3.3 Solution with Trained Basis and No Sparsity Constraint

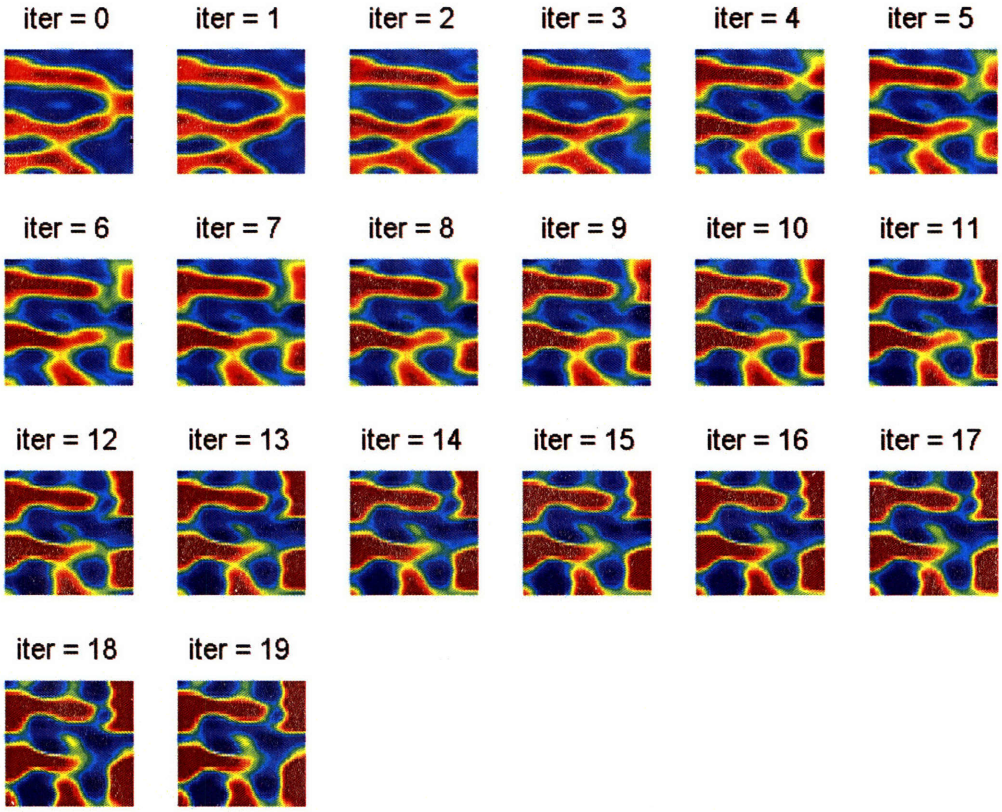
In this example, a similar problem to Section 4.3.1 is considered, except that the training library is used to pre-select more relevant basis vectors for the reconstruction. The solution is expected to depend on the accuracy of the prior information. However, since the prior library is statistically representative of the unknown features in this example, the solution is likely to improve compared to the example of Section 4.3.1.

Figure 4.3 illustrates the results for this example. It is observed that even when prior information provides knowledge about orientation of the existing channels, the algorithm fails to detect the shape and connectivity of the two channels. Unlike the convex examples in Chapter 2, it is possible that the minimization algorithms in this chapter are trapped in a local minimum due to the nonlinearities in the model. The results suggest that, without the sparsity constraint, the problem is still under-constrained even when dynamic measurements are used to further constrain the solution.

4.3.4 Solution with Trained Basis and Sparsity Constraint

The last example of this section evaluates the reconstruction performance when sparsity constraint is included in the objective function and the prior training library is used to select more relevant basis images for the reconstruction. The minimization iterations and final solution are shown in Figure 4.4. As expected, the reconstruction quality is better in this case than all the other examples in previous sections since all sources of information are incorporated into the problem.

(a) Permeability estimates after each minimization iteration



(b) Reduction in objective function after each minimization iteration.

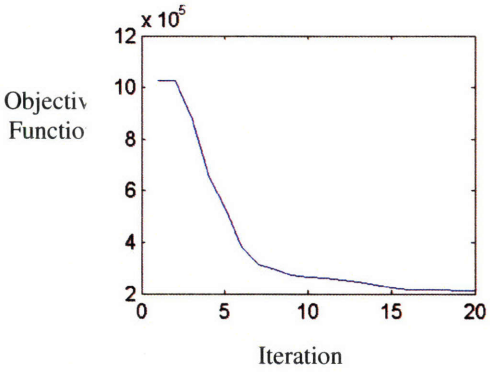
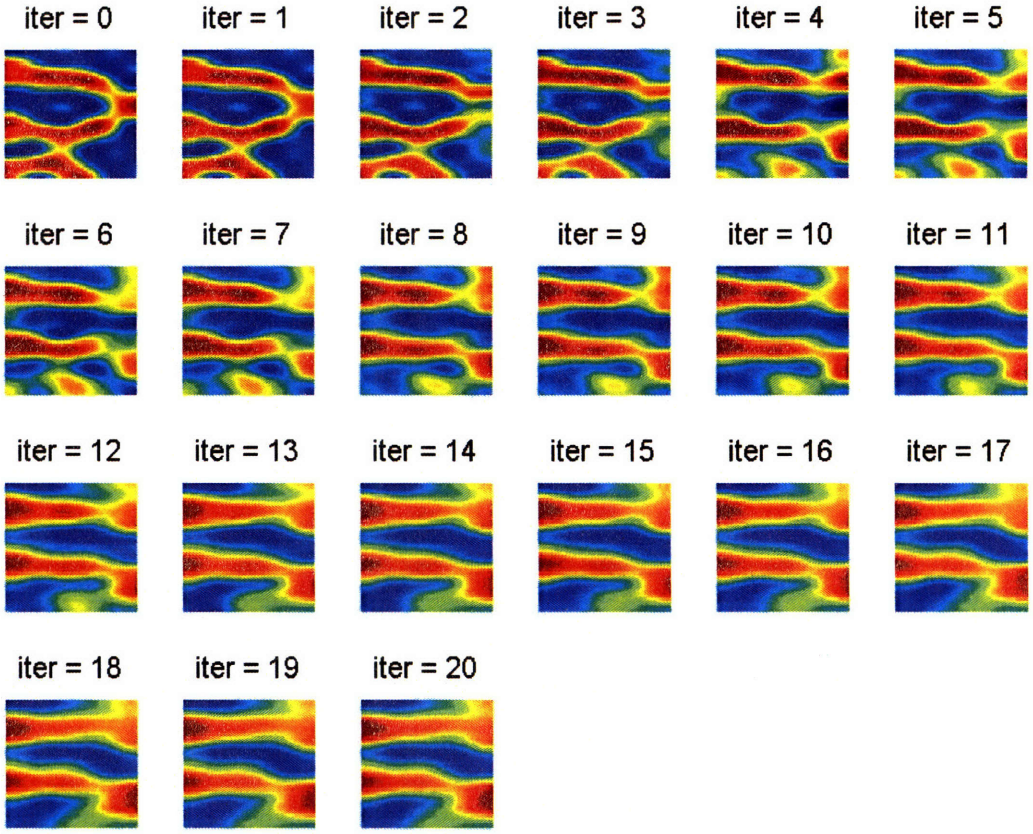


Figure 4.3 Reconstruction results for inversion with static and dynamic measurements using prior information and without sparsity constraint: (a) log-perm after each minimization iteration; (b) reduction in the objective function at each iteration.

Figure 4.4 can be compared with Figures 4.3 and 4.2 to evaluate the effect of adding sparsity constraint and prior model, respectively. It is evident from these comparisons that sparsity has a stronger effect on the quality of the solution than prior training does. This is not unexpected because in these examples the important basis images are present in both trained and untrained cases (from Chapter 2, even S=15 coefficients out of 78 can provide a reasonable approximation). The untrained basis, however, contains some irrelevant basis images (e.g. those that represent vertical variability for instance) that are truncated in the case of the prior trained basis. On the other hand, the sparsity constraint is more important because it removes non-sparse local solutions.

Finally, Figure 4.5 shows saturation predictions using the estimated permeabilities in above experiments. Figure 4.5a has the predictions using the true permeability model and the rows in Figure 4.5b, i.e. (b1)-(b4), contain the same predictions using the final permeability estimates in each of the experiments described above. The predictions with permeability estimates under sparsity constraint, (b2) and (b4), are similar and closer to the true predications. However, the main absent feature in these two predictions is the leftover oil on the top section of the reservoir at the end of the simulation (36 months) as indicated by the true predictions (first row). This appears to be mainly due to the lack of observations from the mid-section of the field to reveal the presence of the low-permeability feature on the top.

(a) Permeability estimates after each minimization iteration



(b) Reduction in objective function after each minimization iteration.

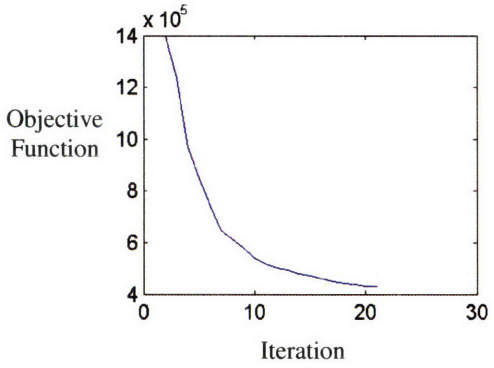


Figure 4.4 Reconstruction results for inversion with static and dynamic measurements without using prior information and sparsity constraint: (a) log-perm at each minimization iteration; (b) reduction in the objective function after each iteration.

On the other hand, the predictions without using the sparsity constraint, (b1) and (b3), deviate greatly from the true predictions as expected from the poor permeability estimation results. The water saturation fronts (blue) in these two cases move faster in the beginning (first 6 months) due to overestimated permeabilities on the left side of the reservoir. However, these saturation fronts are delayed when they hit the discontinuity on the right hand side. It appears from these figures that these two effects tend to cancel each other and the net arrival time of the front at the production side is similar to predictions with the connected true permeability.

Based on the results presented in this section, dynamic measurements appear to provide useful information about unsampled points of the permeability to guide the reconstruction algorithm. When sparsity constraint is not used, the addition of the new dynamic measurements is not sufficient to constrain the inversion and minimization is likely to be stuck in a local solution. However, when sparsity is used to constrain the likely solutions, the solution appears to better resemble the true permeability field. This however, does not mean that a global minimum is found as the objective function is not convex and several local solutions may exist. Nevertheless, the sparsity constraint tends to avoid irrelevant and disconnected local solutions of the ill-posed inverse problem. In a sense, by requiring sparsity in the DCT domain, continuity seems to be achieved in the spatial domain, which is desirable for characterization of geological structures.

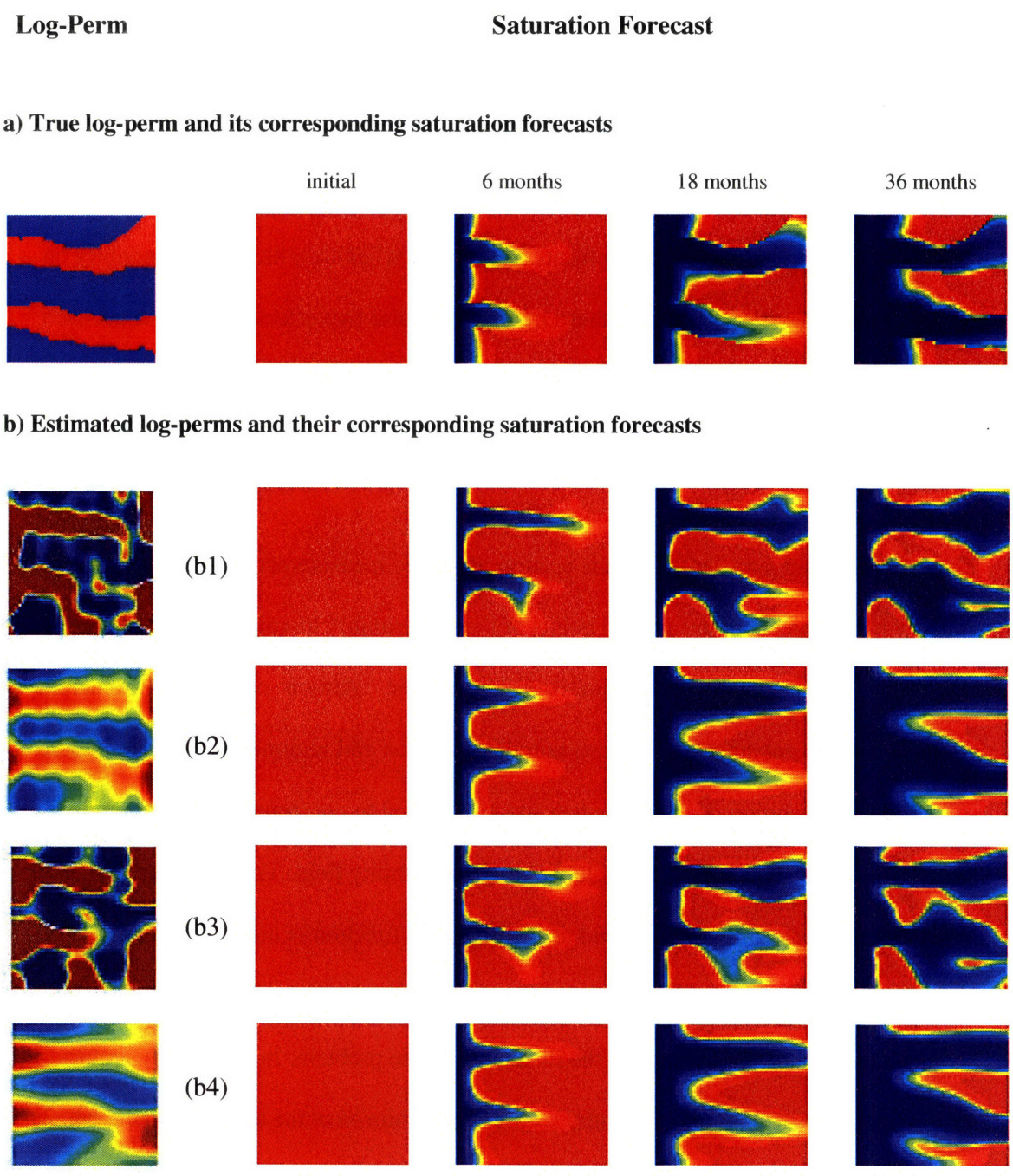


Figure 4.5 Saturation forecasts using the estimated permeabilities for experiments in section 4.3: (a) true log-permeability and saturation distribution; (b1)-(b4) estimated permeabilites and saturations for sections 4.3.1 – 4.3.4, respectively.

Chapter 5

Conclusions and Future Directions

This chapter presents a summary of the conclusions that can be drawn from previous chapters of this thesis. It also outlines the contribution of this thesis and possible future research directions.

5.1 Thesis Conclusion

In this thesis, an estimation approach was introduced for solving ill-posed inverse problems with unknown parameters that are approximately sparse in a transformed domain such as DCT. The formulation has its origin in basis pursuit and compressed sensing, with widespread application in signal processing and data compression. The method was examined in several interpolation examples, starting with a setup that satisfied the theoretical requirements of the formulation. Some of the restrictive assumptions that are hard to realize in practice were relaxed to make the method applicable to realistic problems. The formulation was also considered in dynamic inverse

problems where observations of time-variant quantities are used to constrain the proposed reconstruction algorithm further.

In general, it is concluded that the sparsity constraint can improve the solution of the ill-posed problems, in which the unknown parameters have an approximately sparse representation in compression domains such as DCT. The theoretical perfect reconstruction was observed in the interpolation problems only when a relatively large number of observations (compared to level of sparsity) at random locations were used. Unfortunately, perfect reconstruction is more sensitive to degree of sparsity and number of observation than it is to the dimension of search space. This can have practical implications in the areas where the number of observations and the level of sparsity can not be controlled. Therefore, applicability of the approach in situations where the observations are limited and the sparsity requirements may not be perfectly satisfied was evaluated.

The quality of reconstruction was adversely affected when the observations were fixed and limited in space and the original image was approximately sparse (with many non-zero but small coefficients). However, compared to the experiments in which no sparsity constraints were used, the estimation results were significantly better when sparsity constraint was included.

Several nonlinear inversion experiments with dynamic observations were also conducted to examine the influence of the sparsity constraint in these settings. The results of these

experiments are in agreement with the general conclusions drawn from the interpolation examples; that is, the sparsity constraint enhances the reconstruction of the unknown large scale features (patterns) by providing a systematic way of identifying the significant basis vectors and tuning their corresponding coefficients. The proposed estimation approach seems to be a suitable framework for detecting continuous objects such as geological facies and channels. These results may also hold in other similar applications where the solution to under-constrained inverse problems with (nearly) sparse unknown parameters in an incoherent basis is desired, e.g. in geophysics, medical imaging, and pattern recognition.

5.2 Thesis Contributions and Future Research Directions

The original contributions of this thesis can be summarized as:

1) *Adaptation of the compressed sensing framework for solving linear interpolation problems to estimate smooth and continuous parameters that exhibit "approximately" sparse behavior in an incoherent transform domain such as DCT.*

2) *Extension of the new formulation to nonlinear inverse problems with a dynamical state-space model and more realistic assumptions about observation quantity and distribution.*

While the proposed formulation in this thesis has its origin in compressed sensing theory, the novelty and contribution of this thesis lies in the adaptation and successful application of this theory to problems in which theoretical assumptions about observed quantities and their spatial distributions, as well as the perfect sparsity of the original signal, were violated. Also, reduction of the search subspace dimension (N) was proposed to improve the computational complexity and efficiency of the algorithm. In addition, the reconstruction problem was applied to nonlinear and ill-posed inverse problems with only *approximately* sparse unknown parameters and limited observations in space.

In summary, the results presented in this research suggest that sparsity constraint provides a promising approach for inversion of ill-posed problems with approximately sparse unknowns in a complementary transform domain. However, there are several issues to be studied before successful application of the current form of this method is achieved in practice. The study in this thesis has been limited to simple two dimensional examples with a good prior model (when it was used). Therefore, success in generalizing the proposed approach to realistic situations with more complex three dimensional features cannot be assumed. It will be interesting to see the performance of the proposed algorithm in retrieving complex three dimensional features.

Finally, these results and the important implications that they may have in dealing with many real-world ill-posed inverse problems (such as geophysical inverse problems, subsurface characterization, and medical imaging to name a few) call for evaluation of the proposed inversion algorithm in large-scale problems. Several topics can be studied in

future, including estimation of more complex features and patterns, selection of a more specialized basis than the DCT for a given application, incorporation of remotely sensed data such as 3D and 4D seismic, and integration of observations at different scales and with different resolution.

References

1. Jain A.K. (1989): *Fundamentals of Digital Image Processing*, Prentice Hall.

2. Gonzalez, R.C., Woods, R.E. (2002): *Digital Image Processing*, 2nd ed., Prentice Hall, Upper Saddle River, NJ.

3. Aziz K., Settari A. (1979): *Petroleum Reservoir Simulation.* Applied Science Publishers LTD, London.

4. Mattax, C.C., Dalton, L.R.: *Reservoir simulation SPE monograph*, Volume 13, 60-61.

5. Gavalas G.R., Shah P.C., Seinfeld J.H. (1976): "Reservoir history matching by Bayesian estimation.", *Soc. Petrol. Eng. J.*, 16(6):337–350.

6. Ahmed A., Natarajan T. Rao K.R. (1974): "Discrete cosine transform". *IEEE Trans. Biomed. Eng.* C23: 90-93.

7. Rao K. R., Yip P. (1990): *Discrete Cosine Transform: Algorithms, Advantages, Applications,* Academic Press, Boston.

8. ECLIPSE reservoir simulator (2006): "Manual and Technical Description," *Schlumberger GeoQuest*, Houston.

9. Strebelle S., Journel A.G. (2001): "Reservoir modeling using multiple-point statistics". paper SPE 71324 in the proceedings of the *SPE Annual Technical Conference and Exhibition*, New Orleans.

10. Remy N. (2004): "S-GeMS: A geostatistical earth modeling library and software". Ph.D. thesis, Stanford University.

11. Abdelhalim M. B., Salama A. E. (2003): "Implementation Of 3D-DCT Based Video Encoder/decoder System", *International Symposium on Signals, Circuits and Systems*, 389 – 392.

12. Candes E., Romberg J., and Tao T. (2006): "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inform. Theory, vol. 52,no. 2, pp. 489–509.

13. Donoho D. (2006): "Compressed sensing," IEEE Trans. Inform. Theory, vol. 52, no. 4, pp. 1289–1306.

14. Candès E., Tao T. (2006): "Near optimal signal recovery from random projections: Universal encoding strategies?". *IEEE Trans. on Information Theory*, 52(12), pp. 5406 – 5425.

15. Candès, E., Romberg J. (2005): "Practical signal recovery from random projections". In *SPIE Symposium on Electronic Imaging*.

16. Bloomfield P., Steiger W. (1983): "Least Absolute Deviations: Theory, Applications, and Algorithms". Progr. Probab. Statist. 6, Birkhäuser, Boston, MA.

17. Jafarpour B., McLaughlin D.B. (2007): "Permeability parameterization with the discrete cosine transform". *In Proceedings of SPE Reservoir Simulation Symposium*, Woodland, TX.

18. Tikhonov A. N., Arsenin V. I. (1977): *Solution of Ill-Posed Problems.*Washington, DC: Winston.

19. Schweppe F.C. (1973): *Uncertain Dynamic Systems*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

20. Alliney S., Ruzinsky S. (1994): "An algorithm for the minimization of mixed l_1 and l_2 norms with application to Bayesian estimation". *IEEE transactions on signal processing*, 42(3):618-627.

21. Reiniger R., Gibson J. (1983): "Distribution of the two-dimensional DCT coefficients of images". *IEEE Trans. Comm.*, vol. COM-31(6), pp. 835-839.

22. Eggerton J.D., Srinath M.D. (1986): "Statistical distribution of image DCT coefficients", *Comput. & Electr. Eng.*, vol 12, pp.137-145.

23. Muller E. (1993): "Distribution shape of two-dimensional DCT coefficients of natural images". *Electr. Letters*, vol. 29(22), pp. 1935-6,28th Oct.

24. Eude T., Grisel R., Cherifi H., Debrie R. (1994): "On the distribution of the DCT coefficients'' *Proc. IEEE ICASSP*, Adelaide (Australia), vol. *V*, pp. 365-8.

25. Tarantola A. (2004): Inverse Problem Theory. Methods for Model Parameter Estimation, Philadelphia, PA: SIAM.

26. Ruzinsky S. A. (1989): "Sequential least absolute deviation estimation of autoregressive parameters". *PhD. Thesis*, Illinois Inst. of Technol., Chicago, IL.

27. Chen S.S., Donoho D.L., Saunders M.A. (2001): "Atomic decomposition by basis pursuit". *SIAM Review* **43** (1), pp. 129–159

28. The MathWorks--MATLAB: http://www.mathworks.com/

29. Gelfand I.M., Fomin S.V. (2000): *Calculus of Variations*, Dover Publications.

30. Bennett A.F. (2002): *Inverse modeling of the ocean and atmosphere*, Cambridge University Press, Cambridge.

31. Gill P.E., Murray W., and Wright M.H. (1983): "Practical Optimization", Academic Press, New York.

32. Nocedal J., Wright S.J. (2006): *Numerical Optimization*, Second Edition, Springer.

33. Zhang F., Reynolds A.C. (2002): "Optimization algorithms for automatic history matching of production data.", *8th European Conference on the Mathematics of Oil Recovery*, Freiberg, Germany.

34. Jafarpour B., McLaughlin D.B. (2008): "Estimating channelized reservoir permeabilities with the ensemble Kalman filter: the importance of ensemble design". Accepted for publication in SPE Journal.