UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

IMPROVING ADAPTIVE SENSING SCHEMES IN THE RECOVERY OF *2-D* SIGNALS UNDER SPATIAL CORRELATION AND SPARSITY CONSTRAINTS

TESIS PROPOSAL Doctorado en Ingeniería Eléctrica

FELIPE ANDRÉS SANTIBÁÑEZ LEAL

PROFESOR GUÍA:
PH.D. JORGE SILVA SANCHEZ

MIEMBROS DE LA COMISIÓN:
TO BE DEFINED
TO BE DEFINED

Este trabajo ha sido parcialmente financiado por CONICYT PHD FELLOWSHIP 2013

SANTIAGO DE CHILE
DECEMBER 2015

Abstract

Although *Geostatistic* tools have become standard for characterization of the spatial distribution of geological subsurface structures, the related inference problems for low acquisition levels still pose a complex issue only tractable by simulation tools. Against that background, in the last decade several methods for low-rate sampling and sparse representation were developed providing insights into use of additional prior information to achieve better performance in reconstruction and characterization tasks.

Based on these achievements, the new challenge is to incorporate several tools from the *state of art* in signal processing and stochastic modeling to improve this kind of inference problems. We propose a comprehensive study of inverse problems at low-rate sampling with strong focus in Geosciences and in particular for the reconstruction of binary permeability channels.

For the first part of this work, we will work on theoretical formulation and experimental analysis of the *Optimal Well Placement* (*OWP*) problem. It tries to find the best way of distributing measurements to optimize sensing/locating resources in areas of mining and drilling. This work aims at obtaining the optimal positions for a given amount of available measurements. The characterization of the uncertainty will be a central piece of this formulation. In particular, the *OWP* problem will be addressed from the perspective of minimizing the remaining field uncertainty and sequential algorithms will be proposed to solve it. We conjecture that *OWP* based locations will be distributed on transition zones, and then we will assess this behavior for common sensing schemes.

For the second part of this work, we propose to derive image restoration/reconstruction techniques by using sparse regularization methods such as Noisy Compressed Sensing (Dantzig Selector and *Basis Pursuit DeNoising*). Specifically, we focus on incorporating spatial correlations by experimental covariance matrices and we derive several approaches to simplify scenarios with thousands of unknowns. The Subsurface characterization problem for binary channels will be addressed from sparse representation perspective, incorporating Multi-Point Simulations (*MPS*) and training images as an additional source of spatial dependence.

Finally, we propose the integration of these two topics (adaptive sampling design and inference by Noisy Compressive Sensing) to improve subsurface characterization techniques based on adaptive sampling and sparsity promotion by the use of Adaptive Compressive Sensing approaches.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

In complex real scenarios the task of characterizing an actual underlying model representing a system (for instance, an industrial process or an environment to be studied, prospected or exploited) is extremely hard and surrounded by uncertainty. This task is formally defined as an inverse problem (belonging to the problems classification described in Table 1.1) and rely on the mapping relation among the acquired information (about the actual system), termed data or measurements, and the parameters for the underlying model. The inverse problem reduces to finding the value for these parameters from the knowledge of data, describing an attempt to construct a model coherent with the available evidence [1]. In this context, a prior knowledge about the nature of the required model describes our best effort to characterize the space where the parameters belong.

Table 1.1: Forward and Inverse problems in engineering and science

	Forward Problems		**Inverse Problems**	
	System Design	**Convolution**	**System identification**	**Deconvolution**
Input:	Known	Known	Known	**Unknown**
System:	**To be designed**	Known	**Unknown**	Known
Output:	Predefined	**Unknown**	Known	Known

1.1.1 Sensing and Inverse problems

In inverse problems, the relationship ($g(\cdot)$) between the distribution of required set of parameters $\mathbf{X}$ and the model outputs (measurements) $\mathbf{Y}$ is often described by a complex and non-linear forward model. We can describe the data acquisition process (i.e. *sensing*) and its relationship with the forward model by:

$$\mathbf{Y} = g(\mathbf{X}). \tag{1.1}$$

The problem can be formulated as follows: Given a set of measures Y recreated by a forward model $g(\mathbf{X})$, we want to obtain the set of parameters $\mathbf{X}$. As $\mathbf{X}$ is unknown, we need to define an objective function that estimates the match between Y and $g(\mathbf{X})$. In general, the problem has non-unique solution (ill-posed problem) because there are many possible coherent model parameters.

A classical definition for the objective function is the one described in eq. (1.2) :

$$G(\mathbf{Y}, \mathbf{X}, g(\cdot)) = \|\mathbf{Y} - g(\mathbf{X})\|_p\,, \tag{1.2}$$

where $\|\cdot\|_p$ denotes the p norm.

We can also consider some level of uncertainty by incorporating a noise component in the relationship (1.1), obtaining a model given by:

$$\mathbf{Y} = g(\mathbf{X}) + \nu. \tag{1.3}$$

1.2 Problem Relevance

1.2.1 Inverse Problems and Sampling

In a wide range of applications the model outputs $\mathbf{Y}$ belongs to a high dimensionality space, becoming unfeasible or at least impractical its full observation. Thus, considering $g(\mathbf{X})$ as a signal residing in a high dimensional space $\mathbb{R}^N$, we could only have access to a finite number of measurements, $\mathbf{Y}_{\mathbf{Obs}} \in \mathbb{R}^m$. Therefore, we can measure:

$$\mathbf{Y}_{\mathbf{Obs}} = A(g(\mathbf{X}) + \nu) + \eta, \tag{1.4}$$

where the function $A(\cdot)$ represents the sampling scheme and η is the noise associated to the measurements.

Traditional signal processing theory (Sampling theorem, from section 1.4.2 at [2]) stated that data acquisition systems require to sample signals at a rate exceeding twice the highest spatial/temporal frequency for the purpose of characterizing the band limited signals. It is the principle behind most of imaging acquisition and audio recorders. However, in many practical problems only $m << N$ measurements are accessible to solve the ill-posed problem from Eq. (1.4). Nevertheless, a wide range of these applications requires the recovery of signals that are sparse or at least compressible in an appropriate transform domain. This fact has motivated the adoption of sparsity promoting solutions to solve Eq. (1.4) [3, 4, 5].

1.2.2 Geosciences and the Sampling Problem

Associated with the relationship in Eq. (1.4), we need to stipulate the forward model, $g(\cdot)$. It can either be supplied from a physical model, empirical evidence or from a statistical model that connects the observed data and the model parameters.

An important case of interest in the scope of this thesis proposal corresponds to tasks related with reservoir characterization. In the characterization of a reservoir as the shown in Fig. 1.1, several variables are relevant including discrete ones (such as fluid filling indicators, rock or sediments types), or continuous ones (such as porosity and permeability) [6, 7, 8]. One of the main challenges lie in the fact that usually no direct observations, or just a reduced amount of these, are available leading to the use of indirect data for the inference process.

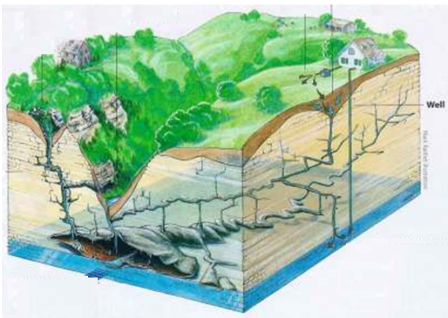


Figure 1.1: Example of a *3-D* geological reservoir.

The limited available measurements for the variables of interest makes any inference attempt very challenging. The aforementioned implies that geophysical problems are usually undetermined and ill-posed. In short, characterizing a reservoir is based on the sources of information: *well observations* (direct samples), production data, and seismic data [8].

Seismic data is usually available at large scale, provided even through the entire reservoir. Nevertheless, related data is sampled on coarser grids and associated with several uncertainties and noise levels.

Well observations generally consist in well logs taken from process as described in Fig. 1.2. *Well data* is only available in areas of measurement provided by the existing wells. However, *wells* are sampled on a fine grid in the paths of interest providing better details than seismic data and the uncertainty associated with these measures is smaller than the seismic sources.

In the case of *production data*, the information is acquired along the production process

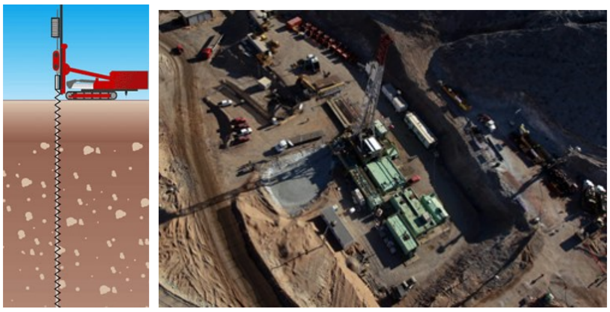


Figure 1.2: Sampling scheme.
Left: Example of a well sampling scheme. Right: Example of an actual well sampling system.

in *production wells*. This kind of data takes into account global factors related with features of reservoir that makes a great difference with the other sensing modalities. *Historical data* considers scenarios with large volume of information collected allowing history matching approaches [6, 8, 9] and moving away from the scope of this thesis proposal.

1.2.3 Geosciences and Uncertainty

Aforementioned, the information sources for reservoir characterization contain uncertainties [10, 11, 12]. For this reason, an accurate description of the spatial distribution of a subsurface model is essential for reservoir characterization, which plays a key role for mineral and fuel exploration and production [8, 10, 11, 13, 14, 15]. Therefore, several techniques have been developed in order to estimate, quantify, and represent these uncertainties [16, 17].

Unfortunately, the access to true observations of the subsurface structures is not possible because direct measurements are limited in number and unevenly distributed. Based on the above, the geological characterization is performed by using several indirect acquisition processes [6, 8]. In this context, a stochastic modeling of the problem is essential. It should also be noted that measurements may be inaccurate making the characterization of reservoirs even a harder problem [16].

Reservoir properties at various grid locations (pixels on a discrete two dimensional representation) are largely unknown, hence each property of interest at every grid block (or pixel) is modeled as a random variable whose variability is described by a probability measure. The reservoir characterization relies not only on reduced portion of available data but also in its placement. This rises the importance of the optimal sensing placement problem and recovery approaches for the inverse problem in Eq. (1.4), which is the focus of this work.

1.3 Problem Statement

1.3.1 Motivation

In this thesis proposal, the focus is on the field characterization of subsurface structures from spatial observations and its relationship with *sampling theory* and *sparse representations*.

Beyond prospecting processes, both exploitation by mineral blasting and short-term mine planning consider and could take advantages of sensing design and inference tools. While blast hole drilling systems has been focused on efficient drilling instead of high precision sampling, short term mine planning use a medium scale sampling to choose mining units based on estimated distribution of ore and waste. As units classified as ore will be sent to the plant while waste units to the waste dump, mistakes in this classification process probably has a significant impact in economic terms.

Design of blast hole drilling and short term mine planning has been stated as structured grid sampling where the degrees of freedom only consider the scale of the grid. In this context our sensing design and inference approaches could improve decision-making issues in mining units classification.

1.3.2 Channelized Structures

This proposal, preliminary, emphasizes on a classical geosciences scenario related with subsurface channels systems. As previously shown in Fig. 1.1, several continuous or discrete variables could be modeled by channelized structures. In addition, a statistical model is required as part of the forward model describing the channelized structures, where non-stationary assumption could be necessary to reproduce channel-like features [8, 18].

A representative variable of interest corresponds to a binary permeability channel (Fig. 1.3). Structural geological models (an essential tool in reservoir characterization studies, exploration and prospecting) are used to describe this kind of fields by map estimation and map making tasks.

Figure 1.3: Binary permeability channel.
Example of a *2-D* representation of a binary channelized permeability field.

Formally, a map is a descriptive numerical representation of the spatial distribution of a subsurface attribute (such as thickness, permeability, porosity, flow rate, etc.) [6, 19, 20]. The general process of the map estimation (i.e. inference of the spatial distribution) is described in Fig. 1.4, where given a reduced number of observations the final goal is to infer the underlying media or at least a related significant feature (such as first and second order statistics, connectivity, or transport metrics).

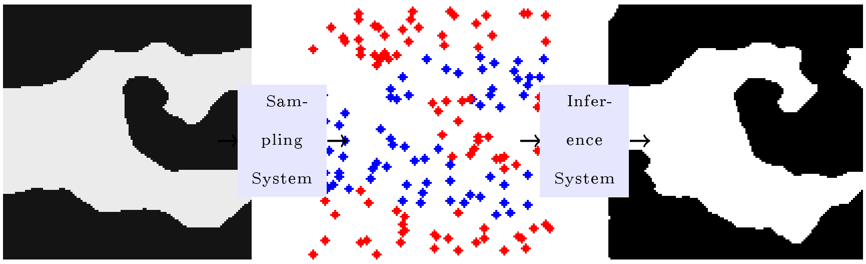


Figure 1.4: Basic inverse problem scheme.
Example of a basic scheme of inverse problem related with *2-D* channelized structures characterization.

1.3.3 Problem Approaches

We can assess the inverse problem of characterizing a field based on few measures through two non exclusive ways. On the one hand, we can ask for the (near)optimal (in some relevant meaning) sensing schemes optimizing $m << N$ observations by some *informative decision* function, guiding the problem to find the best distribution for the observations in order to improve the knowledge of the original signal $\mathbf{X}$. This problem is usually termed as *Optimal Sensor Placement* (*OSP*) [21, 22, 23]. On the other hand, given an appropriate sensing scheme, we can ask for the minimum number of $m << N$ measurements required for a (near)perfect reconstruction/estimation of the wanted signal $\mathbf{X}$ (from ill-posed problem of the eq. 1.4) by taking advantage of desired structured features of the signal (i.e. sparsity, compressibility, spatial or temporal correlations, etc) and the sampling scheme.

In addition, the *Geosciences* community developed several statistics tools in order to achieve good estimations for describing structures with spatial dependence such as channelized fields.

In the next sections we describe the *state-of-art* approaches used in both global characterization and field estimation: *Geostatistics* approaches (sec. 1.4), sensing design based approaches (sec. 1.5) and sparse promoting methods (sec. 1.6).

1.4 Methods used in Geostatistic Analysis

In recent decades, geologists have achieved realistic representations of the internal structure of reservoirs considering complex and heterogeneous geological environments through the use of *Geostatistics*. *Geostatistics* deal with spatially correlated data such as facies, reservoir thickness, porosity, and permeability [8]. *Geostatistics* tools allow to evaluate potential exploration and production zones, where a main issue is to define well locations.

The term *Geostatistics* usually refers to the branch of spatial statistics that is concerned with the analysis of an unobserved spatial phenomenon $X = \{X_{(u,v)} : (u, v) \in D \subset \mathbb{R}^2\}$ (for the *2-D* case), where D denotes a geographical region of interest. When the spatial coordinates are discretized the subset D is comprised by only N positions allowing the representation of the field as the set $X = \{X_{\mathrm{i}} : \mathrm{i} \in \{1, \dots, N\}\}$.

1.4.1 Two-Point Statistics

When large amount of measurements are available, then is possible to use interpolation techniques on the observed data. For example, *Kriging* techniques attempt to achieve interpolation using variograms as shown in Fig. 1.5 that captures the spatial correlation of data with excellent precision [6]. Classical *Kriging* approaches are based on *two-point* statistics [6].

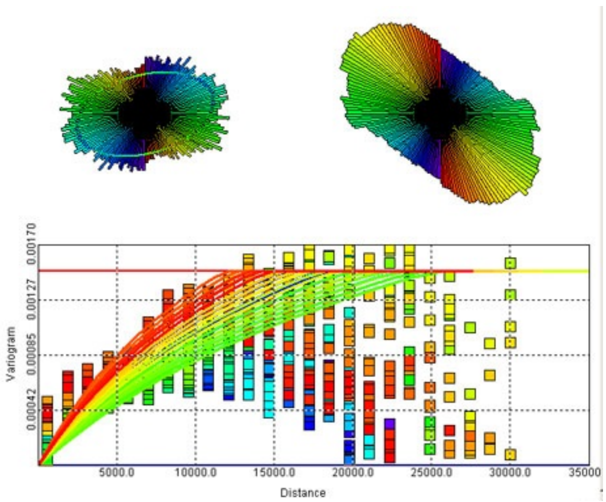


Figure 1.5: Example of a classical variographic analysis.

Geostatistics deals with stochastic processes defined in a region D, with $D \subset \mathbb{R}^{\mathrm{d}}$ and considering d = 1, 2, or 3. For continuous stochastic fields, a classical assumption is that the field X is modeled by a stationary and isotropic *gaussian* process, with zero mean, constant variance σ^2 and autocorrelation function given by $\rho(X_{\mathrm{i}}, X_{\mathrm{i}+h}; \phi) = \rho(\parallel h \parallel; \phi)$, $\forall \{\mathrm{i}\} \in D$.

Traditional *Kriging* tools (based on the autocorrelation function) only takes into account *two-points* statistics [19]. Thus, these techniques based on variograms reproduction tends to fail in the modeling of realistic geological facies[1]. Due to the variographic analysis measures only facies continuities between two regionalized variables positions, variograms fail describing curvilinear or multiscale structures that requires the inference of joint correlations of facies at multiple variables positions. Therefore, these models are unable to properly represent long-range continuities or discontinuities of subsurface fields, misrepresenting the real reservoir connectivity. This translates in poor reservoir performance forecasting [24, 13, 21].

1.4.2 Object-based Simulation

Both geologists and reservoir engineers have a keen interest in local scale details describing reservoir heterogeneities. For this, stochastic simulations provide an appropriate tool with special emphasis in the regime when a small level of data is available [6, 19].

An initial approach for stochastic simulation was based on object-based simulation. This method simulates many spatial variables by the superposition of some predefined model geometry (e.g. discs, sinusoids, manifolds). Predefined geological shapes require to be manually selected by an expert. In contrast to variographic analysis, object-based methods provides realistic facies structures, but the selection and acquisition of an appropriate conditioning data is a critical limitation [25, 26].

1.4.3 Multi-Point Simulations

Geostatistical simulations provide a powerful tool to reproduce more faithfully and realistically the spatial variability with a small numbers of measurements [18, 24]. A simple scheme is shown in Fig. 1.6. Geostatistical simulations lead to reservoir models that can be constrained to geologic, seismic and production data. These models provide appropriate representations for *geological heterogeneities*, and allow the integration of various types of data at different scale and precisions. In any regime of data, expert knowledge is required to validate and to interpret the applicability of the *Geostatistic tools* [8, 15, 16].

[1]rock structures recognized by its composition or fossil content and mapped by these characteristics.

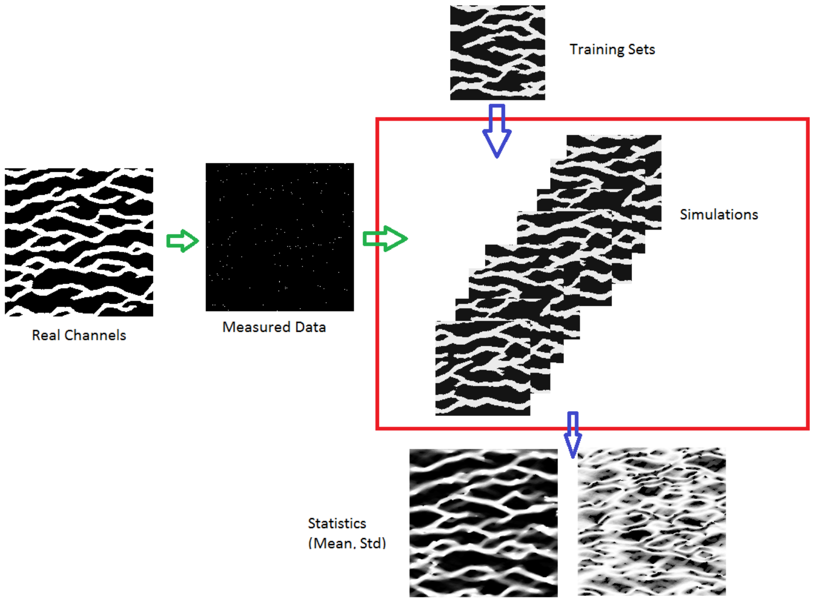


Figure 1.6: Example of a simulation based system.

Deterministic *Kriging* provides good solutions at data conditioning while object-based approaches provides acceptable solutions in the reproduction of geological shapes, but none method is good at both scenarios. Thus, multi-point simulation (*MPS*) was developed to combine the strengths of previously stated approaches. The *MPS* methodology recreates a realistic realization of the target field keeping the flexibility of pixel based simulation methods. At the same time, honoring well or seismic data is imposed by hard data incorporation as the initial state of the simulated media. Furthermore, *MPS* realizations can reproduce complex geological shapes and structures by the estimation of simultaneous statistics at multiple positions from the training image.

MPS and Training Images

Unlike previous methods, *MPS* incorporates a prior model by the use of a training image. It can be defined as a *2-D* or $3-D$ valid realization of a field with the same structures that the target field (e.g. structures like channels, reefs, bars, dikes, oriented facies) representing the full range of possible shapes and its scales. Therefore, training images allow to model complex geological features and their connectivity.

Training image must be considered as a conceptual representation of the geological field instead of an actual realization of it, providing only relative spatial information of the distribution of variables of interest in the field [9].

The generation and selection of training images is an important challenge for the *MPS* methodology. A classical option is the simulation of unconditioned realizations using object-based approaches where an expert user defines the facies shapes and dimensions of interest. In order to select an appropriate training image from a set of available ones, a consistency check has been proposed to compare and validate available well data [7].

MPS was proposed to going beyond *two-point* statistics [27] using training images to describe the full range of structures present in the geological field, but due the limited size of training images only the inference of a very reduced portion of the real *multiple-point* statistics is accurate. Therefore, statistics for large scale structures are usually ignored because can led to undesired discontinuities.

Statistics from Training Images

It is important to note that training images are a source of excepted patterns in the target field. For example, let $X = \{X_\mathrm{i} : \mathrm{i} \in \{1, \dots, N\}\}$ be a categorical field to be simulated, with $z_0, \dots, z_r$ different states defining the alphabet of an individual X_i variable. The *MPS* process is a one pixel at time based approach that works sequentially. In *MPS*, a random path is defined to explore all the field positions to be simulated (excluding hard data positions, if available). Explored positions are then simulated becoming conditioning data for the positions to be explored later in the path sequence.

In a specific unsampled variable X_j, a context based rule is required to define the c closest and most relevant context $X_S = \{X_{S_1} = x_{S_1}, \dots, X_{S_c} = x_{S_c}\}$. The selected variables are chosen from initial hard data and previously simulated variables. Then, the probability that the explored variable X_j has the state z given the conditioning data X_S is estimated by the *Bayes* rule:

$$p(X_j = z | X_S = x_S) = \frac{p(X_j = z, X_S = x_S)}{p(X_S = x_S)}. \tag{1.5}$$

Both $p(X_j = z, X_S = x_S)$ and $p(X_j = z)$ are estimated from an appropriate training image. In particular, $p(X_S = x_S) = \frac{\#(x_S)}{N}$ with N the number of pixels of the training image (in this case equal to the size of the target field to be simulated), and $\#(x_S)$ is the number of occurrences of the specific data pattern x_S in the training image.

The probability $p(X_j = z, X_S = x_S)$ can be estimated as a pattern with one additional conditioning. Then $p(X_j = z, X_S = x_S) = \frac{\#(x_{S,j})}{N}$, with $\#(x_{S,j})$ the number of occurrences of the pattern including the conditioning data and the central variable X_j with the specific value z.

1.5 Methods used in Sensing Design

This approach wants to optimize the selection process of the location for the available measurements. The selection of the best set of informative observations corresponds to a common problem in different contexts such as temperature and light monitoring, sensing contamination in a river, mining exploration, collaborative robotic networks design, statistical experimental design [14, 15, 17, 28]. In general, this optimization problem is NP-hard.

1.5.1 Optimal Well Placement

In our context, the sensing design reduces to the optimal well placement (OWP) problem. In a nutshell, OWP problem addresses the most proper locations to make measurements. The related problem of *optimal sensor placement* has been studied in others scientific areas such as *communications* [23, 29] and *machine learning* [28, 30]. The problem states the optimal (or near optimal) systematic way to take measurements in order maximize the inference system performance or another metric. For example, Fig. 1.7 presents a description of two sensing approaches in a recovery based system.

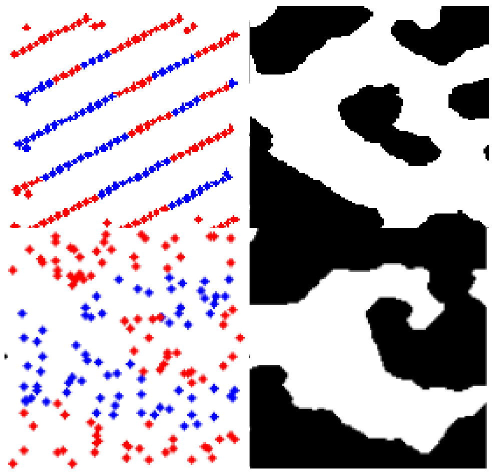


Figure 1.7: Sensing design example and its effect on reconstruction of the field. Upper row: an arbitrary structured sensing scheme (left) and the achieved reconstruction by the measures at this scheme locations (right). Lower row: a near optimal sensing scheme (left) and the achieved reconstruction by the measures at this scheme locations (right). The real image corresponds to the fig. 1.3

Approaches in the literature were oriented to optimize productivity (production functionals) and economic factors [10, 13, 15]. Several optimization methods were proposed to achieve some of these optimal functionals such as adjoint-based gradient [10], simultaneous perturbation stochastic approximation (SPSA) [16], finite difference gradient (FDG) [16], very fast simulated annealing (VFSA) [16], binary genetic algorithm (bGA) [10], continuous or real-valued GA (cGA) [10], and particle swarm optimization (PSO) [10].

While these factors are important in the industrial scenario, the idea of global uncertainty reduction could have a positive impact, directly or indirectly, on the improvement of these other practical factors [14]. In this context, Wellman [12] proposed the use of *information theoretic* tools in the geostatistical analysis for map making tasks. In this specific case, there is a direct connection with *conditional entropy* and its applicability to the characterization of the uncertainty of regionalized variables.

1.5.2 *OWP* and Uncertainty

The *OWP* problem is addressed from view point of uncertainty minimization, where *information theoretic* measures are adopted to formalize the problem using the concept of *a posteriori* conditional entropy.

Considering *2-D* variables with spatial correlation, regionalized variables arises naturally as a suitable model. Here, a regionalized variable X is a square *2-D* random array of variables representing a discrete image of finite size $M \times M = N$ consisting of M^2 discrete random variables $X_{u,v}$.

$$X_{u,v} : (\Omega, \mathbb{P}) \to \mathcal{A} = \{0, \ldots, |\mathcal{A}| - 1\} \quad \forall (u, v) \in \{1, \ldots, M\}^2 \tag{1.6}$$

Given a specific sensing scheme $X_{\text{To Measure}} \subset X$ and assuming that it is possible to estimate the uncertainty for the remaining not sensed field (by the use of *Shannon* entropy $H(X)^2$), then the remaining uncertainty will be updated to $H(X_{\text{Still No Measured}}|X_{\text{To Measure}})$.

A tentative approach is to optimize the location selection of new observations by finding those providing minimal posterior uncertainty for the field X.

1.6 Methods used in Sparse Signal Recovery

In this section we explore some available methods oriented to obtain a reconstruction of a signal from a small amount of measurements under assuming a sparse or compressible model for these signals.

1.6.1 Sparse Signal Recovery

The recovery formulations have a close relation with the inverse problem presented in Eq. (1.4), because the target is the reconstruction of the signal X from the observations Y_{Obs}. In previous sections, signal reconstruction is not addressed because in the regime $m << N$ this is an ill-posed problem.

While the above is correct in general, recent results on *Compressed Sensing* (*CS*) theory provides novel insights for signal reconstruction under the assumptions of desired properties on the signal itself and in the sensing scheme, such as the sparsity at X and the incoherence at the sensing matrix $\mathbf{A}$ [3, 5].

The *CS* theory has found applications on several areas such as signal representation,

[2] $H(X) = \sum_{a \in \mathcal{A}} \mathbb{P}(X = a) \log(\mathbb{P}(X = a))$ when the variables conforming the field X are independent

functional approximation, spectral estimation, cartography, medical imaging, speech signal processing, and sparse channel estimation [5, 31, 32].

1.6.2 Basic formulation

CS attempts to determine the minimal number of observations Y_{obs} required to a stable reconstruction of a sparse[3] signal (with sparsity S). Considering a linear system, in Eq.(1.4) each individual observation is the inner product of the signal X, of size N, with a row vector in the sensing matrix $\mathbf{A}$ as shown in Eq.(1.7).

$$Y_{Obs} = \mathbf{A} \cdot X + \eta. \tag{1.7}$$

The simplest approach for the noiseless case correspond to solve the ℓ_0 minimization problem:

$$(P_0) \qquad \min_{X \in \mathbb{R}^N} \|X\|_0 \text{ subject to } Y_{Obs} = \mathbf{A} \cdot X. \tag{1.8}$$

For the noisy case:

$$(P_0n) \qquad \min_{X \in \mathbb{R}^N} \|X\|_0 \text{ subject to } \|Y_{Obs} - \mathbf{A} \cdot X\|_2 \leq \eta. \tag{1.9}$$

The norm $\|\cdot\|_0$ counts the non-zero entries of the signal X while the norm $\|\cdot\|_2$ denotes *euclidean* norm.

The optimization problems formulated in Eq.(1.8) and Eq.(1.9) are *NP hard*. Due this fact, two main types of methods have been proposed in the last decade to achieve a practical solution.

In the one hand, greedy algorithms like *(Orthogonal) Matching Pursuit* (*(O)MP*) or *Thresholding* methods perform approximations to obtain a suboptimal solution. *Thresholding* method estimates the inner products of the target signal X with all sensing atoms[4] finding the largest ones (in the absolute values) and finally calculating the orthogonal projection onto the span of the selected atoms. The *OMP* method is a sequential approach that selects the most representative atom (i.e. the one with largest absolute inner product with the signal residual) and estimates the signal by updating the residual (the misrepresented part of the observations) by the cumulative selection of atoms. See more details in Table 1.2.

[3] A signal is k-sparse in the canonical domain, if it has at most k terms different than zero

[4] Row vectors conforming the sensing system in the linear model in Eq.(1.7)

Table 1.2: Summary of classical greedy approaches

Goal: An approximated estimation of X from $Y_{Obs} = \mathbf{A} \cdot X$. A_j denotes columns of $\mathbf{A}$ and $\mathbf{A}_\Lambda^\dagger$ the pseudo-inverse of $\mathbf{A}_\Lambda$

OMP	**Thresholding**
initialize: $R = Y_{Obs}$, $\Lambda = \emptyset$ find: $k = \arg\max_j \lvert\langle r, A_j\rangle\rvert$ update: $\Lambda = \Lambda \cup \{\mathrm{i}\}$, $R = Y_{Obs} - \mathbf{A}_\Lambda \mathbf{A}_\Lambda^\dagger Y_{Obs}$ iterate until R based stopping criterion output: $X = \mathbf{A}_\Lambda^\dagger Y_{Obs}$	find: The collection Λ of indices providing largest sparse representation by $\lvert\langle Y_{Obs}, A_j\rangle\rvert$ output: $X = \mathbf{A}_\Lambda^\dagger \cdot Y_{Obs}$

On the other hand, another alternative to address the (P_0) and $(P_0 n)$ is by a convex relaxation of the ℓ_0 norm. For the noiseless case, we obtain the relaxed problem termed Basis Pursuit (*BP*):

$$(P_1) \qquad \min_{X \in \mathbb{R}^N} \|X\|_1 \text{ subject to } Y_{Obs} = \mathbf{A} \cdot X. \tag{1.10}$$

The noisy version of (P_1) is described by the expression termed *Basis Pursuit Denoising* (*BPDN*):

$$(P_1 n) \qquad \min_{X \in \mathbb{R}^N} \|X\|_1 \text{ subject to } \|Y_{Obs} - \mathbf{A} \cdot X\|_2 < \eta. \tag{1.11}$$

Here, ℓ_1-norm is calculated as $\|X\|_1 = \sum |X_\mathrm{i}|$.

The main goal of the *CS* theory is to find guarantees of (near)perfect reconstruction and the associated sufficient conditions. Revolutionary results have established that under certain conditions greedy and *BP* approaches achieves guarantees for perfect reconstruction [3, 33]. For *BPDN* is required that the sensing matrix $\mathbf{A}$ obeys a uniform uncertainty principle, which refers to the presence of well-conditioned submatrices in $\mathbf{A}$.

Formally, let $\Lambda \subset \{1, \ldots, N\}$ be a collection of indices and the $\mathbf{A}_\Lambda$ a submatrix of $\mathbf{A}$ constructed using the columns of $\mathbf{A}$ indexed by Λ. Then, the local isometry constant $\delta_\Lambda = \delta_\Lambda(\mathbf{A})$ is defined as the smallest value satisfying Eq. (1.12) for all vectors X supported on Λ.

$$(1 - \delta_\Lambda)\|X\|_2^2 \le \|\mathbf{A}_\Lambda X\|_2^2 \le (1 + \delta_\Lambda)\|X\|_2^2 \tag{1.12}$$

Finally, the (global) restricted isometry constant is defined by:

$$\delta_S = \delta_S(\mathbf{A}) := \sup_{|\Lambda| = S} \delta_\Lambda(\mathbf{A}), \quad S \in \mathbb{N}. \tag{1.13}$$

If $\mathbf{A}$ has a small restricted isometry constant, i.e. $\delta_S(\mathbf{A}) \le 1/2$, then $\mathbf{A}$ satisfies a uniform uncertainty principle [3].

An important theorem for *BPDN* states that if $\mathbf{A}$ satisfies $\delta_{3S}(\mathbf{A}) + 3\delta_{4S}(\mathbf{A}) < 2$ for some $S \in \mathbb{N}$, and the noisy observations accomplish $Y_{Obs} = \mathbf{A} \cdot X + \xi$ for some $\|\xi\|_2 \leq \eta$, then the signal $X^{\#}$ (solution of the problem $(P_1 n)$) satisfies Eq. (1.14) for an appropriate value of C which only depends on δ_{3S} and δ_{4S} constants [5].

$$\|x^{\#} - x\|_2 \leq C\eta. \tag{1.14}$$

In fact, if $\delta_{4S} \leq 1/3$ then $C \leq 15.41$. In addition, the formulation related with the Eq. (1.14) ensures exact reconstruction for the noiseless problem (P_1) with $\eta = 0$.

While these strong guaranties are only applicable for matrices satisfying the required principle, in practical cases few sensing schemes allow to reach perfect reconstruction. However, *CS* literature remarks that with high probability a $m \times N$ random matrix where the columns are drawn from distributions with certain concentration properties (such as *gaussian* distributions), would have small restricted isometry constants δ_S when $m = \mathcal{O}(S \log(N/S))$.

1.6.3 General Approach for Sparse Reconstruction

Conceptually, *CS* methodology poses a problem that arises quite naturally. The underlying idea corresponds to a regularized problem where the goal is to honor the available observations Y_{Obs} while trying to find the simplest solution. In $(P_0 n)$ honoring of observations is carried out by bounding the *Euclidean* error while the simplicity of the solution is achieved by minimizing ℓ_0-norm or ℓ_1-norm.

It is possible to propose a more general scheme for this regularization approach. The compromise between fitting observations and simplicity can be subsumed by the following unconstrained convex minimization problem.

$$(P_G) \min_{X \in \mathbb{R}^N} \left\| \mathbf{C}_u^{\frac{-1}{2}} \cdot (Y_{Obs} - \mathbf{A} \cdot X) \right\|_p + \gamma \cdot \|\mathbf{W} \cdot X\|_q . \tag{1.15}$$

This regularization problem promotes simplicity using ℓ_q-norm on the target signal X. In addition, a weighting matrix W can be used to incorporate some prior information about the preponderance of some entries of X.

Honoring the observations can be achieved by the ℓ_p-norm of the observation error $Y_{Obs} - \mathbf{A} \cdot X$. The use of the matrix $\mathbf{C}_u^{\frac{-1}{2}}$ allows working with signals that have some level of correlation. Approaches that consider the noisy scenario are usually termed as Noisy Compressive Sensing (*NCS*) approaches.

The problem $(P_1 n)$, known as Least-absolute shrinkage and selection operator ($LASSO$), is a version of (P_G) by ℓ_1-norm regularization with a quadratic constraint [33].

$$\min_{X \in \mathbb{R}^N} \quad \gamma \left\| X \right\|_1 + \frac{1}{2} \cdot \left\| \mathbf{A} \cdot X - Y_{Obs} \right\|_2^2 . \tag{1.16}$$

An alternative approach called *Dantzig Selector* searchs the minimum ℓ_1-norm but with bounded residual correlation [34], as can be seen in the following expression:

$$\min_{X \in \mathbb{R}^N} \left\| X \right\|_1 \text{ subject to } \left\| \mathbf{A}^{\dagger} \cdot (\mathbf{A} \cdot X - Y_{Obs}) \right\|_{\infty} \leq \varepsilon . \tag{1.17}$$

If the underlying target signal is a *2-D* image, an alternate recovery approach is promoting sparsity on gradients of the signal instead of the signal itself. The *2-D* gradient can be calculated from *total variations* (TV), where a possible definition is given by:

$$TV(X) := \sum \sqrt{(D_{h;\mathrm{i}j} \cdot X)^2 + (D_{v;\mathrm{i}j} \cdot X)^2}_{\mathrm{i}j} = \sum \left\| D_{\mathrm{i}j} \cdot X \right\|_{2\mathrm{i}j} . \tag{1.18}$$

Where $D_{h;\mathrm{i}j}$ denotes the gradients on horizontal orientation and $D_{v;\mathrm{i}j}$ the vertical ones. Then, the total variations based $BPDN$ (TV-$BPDN$) approaches is defined by:

$$\min_{X \in \mathbb{R}^N} \quad TV(X) \qquad \text{subject to } \left\| Y_{Obs} - \mathbf{A} \cdot X \right\|_2 < \eta . \tag{1.19}$$

Finally, the *TV-Dantzig Selector* regularization problem can be written as:

$$\min_{X \in \mathbb{R}^N} \quad TV(X) \qquad \text{subject to } \left\| \mathbf{A}^{\dagger} \cdot (\mathbf{A} \cdot X - Y_{Obs}) \right\|_{\infty} \leq \varepsilon \tag{1.20}$$

Although these combinations of regularization factors do not have known theoretical guarantees (as the case of CS), in practical applications of image processing they have demonstrated good performances [35].

Chapter 2

Research Proposal

The *Geosciences* global outcome correspond to the characterization of an unknown complex reservoir field. However, on account of the high number of structures, facies and properties in the subsurface, *geologists* study the field through the description of its individual components.

Thus, we focus on the recovery of a single subsurface property described by a *2-D* regularized variable sampled in the spatial domain. The characterization of these kinds of structures is relevant for mineral prospecting, mining planning, and production stages where the number of available observations is severely restricted by technical and economic factors. Based on the above, the integration of adaptive sensing schemes and sparse promoting approaches is a promising solution to improve classical geostatistical analysis.

2.1 Hypotheses

First, we summarize the main questions related with the channelized signal recovery in low acquisition regimes.

2.1.1 Questions

- What is the minimal observations amount that allows us to achieve an appropriate field characterization?
  - For global characterization, *Geosciences* two-points statistics approaches are useful for high sampling regimes.
  - For full characterization (i.e. description of local details), both sensing design tools and recovery methods based on *sparsity promoting* solvers permit a near perfect reconstruction from low acquisition regimes.

- Given K available measures, what is the *best* location for each one?
  - Sparse promoting approaches provide near perfect reconstructions of sparse signals under specific random sensing schemes.
  - Sampling design theory state that these locations are defined by field complexity and spatial statistics.

2.1.2 Hypotheses

The main hypotheses for this proposal are the following:

- The incorporation of prior information in the design of sampling schemes improves the performance of classical geosciences approaches at low acquisition regimes.
- A specific prior information can be used from different ways involving the achievement of better outcomes by its simultaneous incorporation in sensing design and sparse recovery tasks.
- Prior information based on multipoint simulations (*MPS*) conditioned to hard data (sampled data) allows us to estimate the spatial correlation of binary regionalized variables, and the implementation of adaptive sampling schemes.
- The estimation of the spatial correlation and the incorporation of additional constrains on spatial dependence allows us to approximate covariance keeping the problem computationally tractable.
- Training images provide statistical estimation (i.e. empirical *pdf*s) by pattern occurrences analysis under the assumption of stationarity. At non-stationary scenarios, *MPS* works as an optional way to access to the *pdf*s estimation.
- Adaptive Compressive Sensing (*ACS*) allows us to integrate sensing design and sparse recovery tools together to improve reconstruction algorithms.

2.1.3 Main Objective

The main objective for this research proposal is the development of methods to assist the reconstruction of images describing *2-D* binary regionalized variables by the use of *sensing design* tools, *sparse promoting* techniques and side information from spatial correlation analysis. First, we take advantage of information theoretical tools for the formulation and implementation of adaptive sensing schemes. We focus on regionalized variables with several spatial dependence assumptions by taking advantage of spatial structure and other side knowledge of interest media. Second, given several sensing schemes, we evaluate the sparse promoting techniques addressing practical limitations related with the problem complexity. In addition, a comprehensive study of the integration of both adaptive sampling and recovery methods is proposed by adaptive compressive sensing theory.

2.1.4 Specific Objectives

The specific objectives of this research proposal are:

- Formalize a theoretical framework for *2-D* regionalized variables imposing several spatial dependence constrains in the image model.
- Develop an adaptive sensing design framework using *joint entropy* and *mutual information* to measure uncertainty and spatial structure.
- Study a family of Markov random field models to describe spatial correlation on finite alphabet regionalized random variables.
- Study empirical statistic sources (*MPS* and training images) to estimate the spatial dependence.
- Incorporate a noisy sparse promoting recovery framework by the implementation of *NCS* principles oriented to evaluate its effect on the sampling schemes provided by *OWP*.
- Incorporate virtual noisy measurements in the proposed sparse recovery framework using *MPS* realizations and its empirical covariance as a virtual observations source.
- Integrate the concepts of sensing design and sparse representations on an adaptive compressive sensing approach for *2-D* binary permeability channels reconstruction.

2.2 Methodology

This work will be mainly based on the following steps: data base generation and consolidation applied to channelized structures in geosciences [8, 18, 24, 36], spatial correlation estimation for *NCS* [34, 33], formulation of sampling design problem [23, 13, 14, 11], incorporation of noisy sparse promoting solvers [3, 35, 37] and integration of the sensing and recovery proposed frameworks.

At next sections we describe the proposed methodology required to achieve the objectives posted on sections 2.1.3 and 2.1.4.

2.3 Data Base Definition

2.3.1 Data Base Generation

Here, we identify the models and phenomena of interest to be studied in this proposal. Three kinds of binary channels are proposed as shown in fig. 2.1. These channelized structures are obtained by unconstrained simulations using the geostatistical software *SGeMS* by the *SNESim* algorithm [18, 36].

Figure 2.1: Types of channels proposed for experimental analysis on this thesis.
From left to right: Single channel example (SC1), Multi channel 1 example (MC1), Multi channel 2 example (MC2)

2.3.2 Resizing Data Base

To build the database of permeability channels we consider *2-D* images of size $200x200$. It is an important dimension of the problem setting because the proposed approaches involve optimization problems where computational restrictions need to be imposed.

2.4 Estimation of Spatial Correlations

As the subsurface models of interest (i.e. channelized structures) has spatial correlations, regionalized random fields provide an appropriate framework to capture these spatial dependencies of our models [22]. In this context, we propose to study the spatial distribution as a function of *information theoretic* measures. These *information theoretic* measures will be adapted as a formal criteria to support decision making for *OWP* problem and to describe the remaining uncertainty after the measurements. This work relies on recent contributions [12] concerning the use of information measures for the analysis of geological random functions in terms of uncertainty and spatial dependencies.

We propose the use of statistical analysis from *MPS* realizations as an estimation of field uncertainty under spatial dependence assumption. As exposed in fig. 2.2, we use initial samples from the realization field (a selected realization image from the database) and one previously obtained training image to induce some coherent patterns structure in the simulations[1].

Then, given a set of conditioned simulations we estimate the mean and variance of the individual random variables on the channelized field from it set.

[1]the importance and selection of training images was described in section 1.4.3

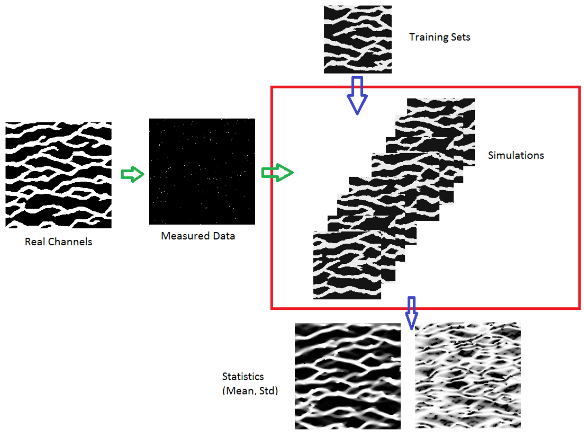


Figure 2.2: Scheme of the estimation of field statistics by the use of some hard data measurements and conditioned simulations using *MPS*.

2.5 Formulation of the Optimal Well Placement Problem

As discussed above in section 1.5, and let K be the number of possible measurements, then the main question that we want to solve is: where do you place these measurements?. In a preliminary approach, we state our problem through stochastic theory and *information theoretic* tools.

2.5.1 Joint Entropy Optimization for OWP

First, we formalize the problem considering the media of interest as a simplified *2-D* finite regionalized variable by the use of a finite alphabet image representing the subsurface channels.

The object to be characterized is a random image (or random field) denoting the subsurface distribution by a collection of finite alphabet random variables $X = \{X_\text{i} : \text{i} \in [N]\}$ where $[N] = \{1, \ldots, N\}$. For every position i in the array, X_i is a random variable with values in a finite alphabet $A^{|X_\text{i}|}$.

Then we can define the collection X_I as the subset of X_i variables with $\text{i} \in I$, where I represent any subset of $[N]$, $X_I = \{X_\text{i} : \text{i} \in I\}$. In addition, the object X^I is defined as the complement of X_I over the collection X, that is to say, $X^I = \{X_\text{i} : \text{i} \in [N] \setminus I\}$.

The probability density function *pdf* of X_i is denoted by $\mathbb{P}_{X_\text{i}}$ in $\mathcal{A}^{X_\text{i}}$, the collection X: $\{X_\text{i} : \text{i} \in [N]\}$ is equipped with its joint probability distribution that we denote by $\mathbb{P}_X$ in $\mathcal{A}^N$. As a short hand, X y $\mathbb{P}_X$ denote the random field and its joint probability, respectively.

Thus, the problem of *OWP* can be posted as the problem of selecting a subset of K elements of $[N]$. Let $\mathbf{F}_K \equiv \{f : \{1, \ldots, K\} \rightarrow [N]\}$ be the collection of functions that select K-elements from N candidates, where every $f \in \mathbf{F}_K$ is a measurement placement rule that models the process of measuring the positions $f(1), f(2), \ldots, f(K)$ in the random field.

Adopting the concept of entropy as a measure of uncertainty of a random variable [38], we propose an algorithm that finds the placement rule f through optimal reduction of *a posteriori* entropy. The criteria used in Eq. (2.1) states that the measurement of the most uncertainty set of K positions will provide an optimal global reduction of the uncertainty for the media of interest (from the point of view of *information theory*).

$$X_f^* = \arg\max_{X_f \subset X} H(X_f) \tag{2.1}$$

More precisely, we try to characterize the conditional posterior entropy that in this context can be expressed as the joint entropy of the entire process minus the joint entropy of the variables measured by f, as shown in Eq.(2.2).

$$H(X^f | X_f) = H(X) - H(X_f) \tag{2.2}$$

Note that objective function in Eq. (2.1) for the search algorithm can use either the left $(H(X^f|X_f))$ or right $(H(X) - H(X_f))$ side of Eq.(2.2).

2.5.2 Formulation and Algorithmic Solution for the OWP

Over the collection of decision rules $\mathbf{F}_K$, we propose to chose the rule that minimizes, in average, the remaining uncertainty after taking the measurements, or the uncertainty of the remaining variables conditioned by the measured variables. More precisely, given a sensor placement decision rule $f \in \mathbf{F}_K$ let us denote by:

$$X_f \equiv (X_{f(1)}, X_{f(2)}, \ldots, X_{f(K)}) \tag{2.3}$$

the measured random vector and by,

$$X^f \equiv (X_{\mathrm{i}} : \mathrm{i} \in [N] \setminus f) \tag{2.4}$$

the non-measured random vector. Note that f subset denotes all the spatial points measured by the *OWP* rule and $f^c \equiv [N] \setminus f$ its complement. Then considering a specific measure $X_f = (X_{f(1)}, \ldots, X_{f(K)}) = x_f \in \mathcal{A}^K$, the remaining uncertainty can be quantify by the *Shannon entropy* [38] of X^f given X_f, *i.e.* $H(X^f | X_f = x_f)$. Note that $H(X^f | X_f = x_f)$ represents the uncertainty conditioning to the specific measured values

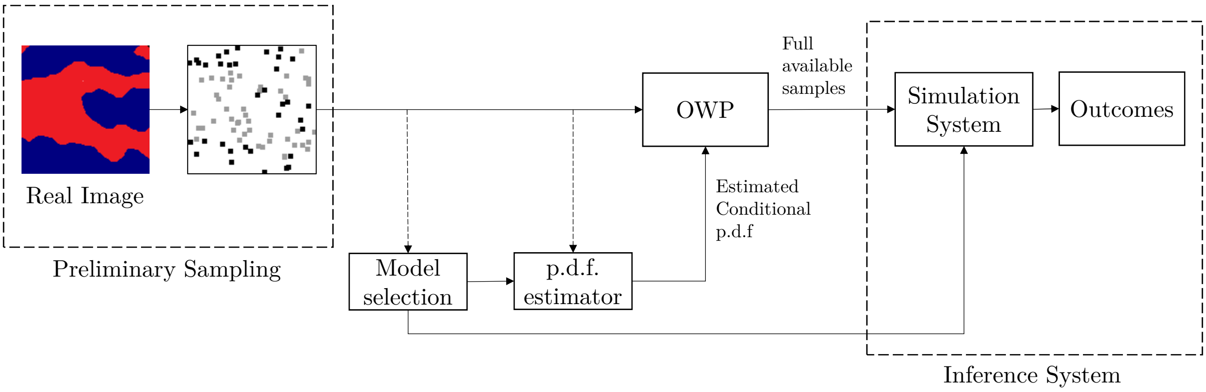


Figure 2.3: General scheme used

$(x_{f(1)}, \ldots, x_{f(K)})$. In practice, we do not have access to this measurements while making a decision on $\mathbf{F}_K$. Consequently, our objective function should consider the posterior uncertainty in average with respect to the statistics of X_f. In other words, we consider the *Shannon conditional entropy* [38] of X^f given X_f, *i.e.*:

$$H(X^f|X_f) = -\sum_{x_f \in A^K} P_{X_f}(x_f) H(X^f|X_f = x_f) \tag{2.5}$$

as the objective function. Then the *OWP* of K-measurements reduces to:

$$f_K^* \equiv \arg\min_{f \in \mathbf{F}_K} H(X^f|X_f) \tag{2.6}$$

which is the solution that minimizes the posterior uncertainty.

2.5.3 Estimation of the Stochastic Field Model

The proposed *OWP* formulation is based on the knowledge of the statistics for the random regionalized variables. In practice, this theoretical model is not available and we need to find the way to estimate this model from empirical data. Sources of empirical statistics are reference images provided by experts, historical information or imposed models. In this work the use of training images and geostatistical simulation tools is proposed as a source for estimating the joint and conditional distributions of X.

2.5.4 Scheme of Sampling Design and Experimental Validation

A basic scheme of the inference process including preliminary blind sampling and adaptive *OWP* is illustrated in fig. 2.3.

The implementation of the different blocks of the proposed system at fig. 2.3 will lead to several practical variants to take advantage of available external information sources. In order to evaluate the *OWP* solutions a performance comparison with random and structured sampling schemes will be performed.

2.6 Formulation and Implementation of Noisy Sparse Promoting Solvers

Given sampling schemes provided by *OWP*, we propose its comparison with some sparse promoting oriented schemes by the modification of classical $L1$ minimizers solver (available in the $L1$ magic and *CVX* software). Our contribution focus on formulating the signal recovery process as a generalized sampling problem. The idea is to take advantage of the sparse nature of channelized structures, the reduced spatial variations of the *2-D* images in the proposed database and the use of a prior statistical model as an additional source of information related with spatial dependencies and pixel uncertainty.

Therefore, we apply principles from Noisy compressive Sensing *NCS* to the reconstruction of binary channels of permeability. For the estimation of the model, we propose the use of *MPS* as a source of statistical data. In particular, it allows the estimation of variance and covariance of the target regionalized variables.

2.6.1 *NCS* and *Whitening* process

We incorporate the noise associated to measurements in the sparse recovery solver relaxing the search space. In this way, given a small amount of noisy measurements from geologic data (m measures from a field of N variables, with $m << N$), the target is to reconstruct the *real* channel. While traditional *CS* approaches deal with time-invariant sparse signals without error in measurements, our motivation was supported by the next hypothesis: 1) *NCS* provides a theory of signal recovery from highly incomplete noisy information, and 2) *MPS* could provide information about signal variability required to the noise characterization on *NCS* framework.

Thus, we implemented and validated a preliminary framework of *NCS* oriented to improve *MPS* performance. Standard *CS* implementation for channelized binary structures was proposed and implemented for others members of the *IDS* Lab [2]. The reader is referred to [39] for more details on this approximation. Classical conditions and theory of *NCS* was described in sections 1.6.2 and 3.4.2.

Here, we extend previous works in order to consider noisy measures and provided a framework to noise characterization. In order to apply the theory of *NCS* we requires a

[2]Information and Decision Systems Laboratory, *IDS* Lab, at Electrical Engineering Department, Universidad de Chile.

signal model with white noise, then a whitening process is required to use the proposed methods. Thus, we consider the next *sensing model* for our problem:

$$I = \Phi \cdot Z \tag{2.7}$$

$$X = I(:) \tag{2.8}$$

$$Y = A \cdot X + \xi \tag{2.9}$$

with

- Z (M x M) : Signal in transformed domain (in our case a *2-D DCT* coefficients matrix of size 200 x 200)
- Φ (M x M) : Transform matrix (inverse DCT in this work)
- I (M x M) : Image in canonical domain
- X (N x N) : Vectorization of signal in image domain ($N = M$ x M, in our case 40000)
- A (m x N) : Sampling matrix (m vectors randomly taken from N x N Identity matrix)
- Y (m x 1) : vector of m measurements (including Hard and Soft Data)
- ξ (m x 1) : vector of noise in measurements, with covariance matrix C_v and mean ξ_{mean}.

At this point the eq. (2.9) only differs from [39] approach in the incorporation of noise component. Here, the noise ξ would be a spatially correlated noise, then we required a *whitening* pre-processing.

A signal model with zero mean noise was described by the subtraction of the mean of noise ξ_{mean} obtaining:

$$Y - \xi_{mean} = A \cdot X + \xi - \xi_{mean} \tag{2.10}$$

Defining zero mean variables and rewriting we obtain the expression in eq. (2.11):

$$Y_0 = A \cdot X + \xi_0 \tag{2.11}$$

From eq. (2.11) required an additional process to obtain a model with a non correlated noise. In order to achieve a sensing model under white noise and assuming the existence of an invertible covariance matrix for ξ_0 we achieved the next formulation:

$$C_v^{\frac{-1}{2}} \cdot Y_0 = C_v^{\frac{-1}{2}} \cdot A \cdot X + C_v^{\frac{-1}{2}} \cdot \xi_0 \tag{2.12}$$

$$C_v^{\frac{-1}{2}} \cdot Y_0 = C_v^{\frac{-1}{2}} \cdot A \cdot X + \eta \tag{2.13}$$

$$\widehat{Y} = C_v^{\frac{-1}{2}} \cdot A \cdot X + \eta \tag{2.14}$$

Replacing the vectorization process we finally obtain the next model:

$$\widehat{Y} = C_v^{\frac{-1}{2}} \cdot A \cdot VEC(\Phi \cdot Z) + \eta \tag{2.15}$$

The eq. (2.15) fits the classical framework of *NCS* for signals under white noise model. The selection of the sampling matrix A satisfies isotropic property, the vectorization process $VEC(\cdot)$ retain spatial dependence in the regionalized field, the basis DCT provides a domain where the signal is compressible, and C_v is estimated as the experimental covariance of realizations of *MPS*.

2.7 Performance Metrics

The performance metrics consider field reconstruction quality analysis and variability of conditioned simulations.

For reconstruction, comparisons between the original image and the reconstructions will be performed. Classical indicators assess the quality of a *2-D* image, X, by comparing it with a reference *2-D* image, Xr, [40] with both images of size M_u by M_v.

The signal to noise ratio *SNR* expressed in decibels *dB*:

$$\mathbf{SNR} = 10 \cdot log_{10} \left(\frac{\sum_{u=1}^{M_u} \sum_{v=1}^{M_v} |Xr_{u,v}|^2}{\sum_{u=1}^{M_u} \sum_{v=1}^{M_v} |Xr_{u,v} - X_{u,v}|^2} \right) \tag{2.16}$$

The peak signal to noise ratio *PSNR* expressed in decibels *dB*:

$$\mathbf{PSNR} = 10 \cdot log_{10} \left(\frac{max(Xr_{u,v})^2}{\frac{1}{M_u \cdot M_v} \cdot \sum_{u=1}^{M_u} \sum_{v=1}^{M_v} |Xr_{u,v} - X_{u,v}|^2} \right) \tag{2.17}$$

The root mean square error *RMSE*:

$$\mathbf{RMSE} = \sqrt{\frac{1}{M_u \cdot M_v} \cdot \sum_{u=1}^{M_u} \sum_{v=1}^{M_v} |Xr_{u,v} - X_{u,v}|^2} \tag{2.18}$$

The mean absolute error MAE:

$$\mathbf{MAE} = \frac{1}{M_u \cdot M_v} \cdot \sum_{u=1}^{M_u} \sum_{v=1}^{M_v} |Xr_{u,v} - X_{u,v}| \tag{2.19}$$

On the other hand, the structural similarity index $SSIM$ tries to estimate the similarity between two images by emulating the human perception [41]. A simplified version of this indicator is given by:

$$\mathbf{SSIM}(X, Xr) = \frac{(2\mu_X \mu_{Xr} + c_1)(2\sigma_{X,Xr} + c_2)}{({\mu_X}^2 + \mu_{Xr}^2 + c_1)(\sigma_X^2 + \sigma_{Xr}^2 + c_2)} \tag{2.20}$$

Here, μ_X (μ_{Xr}) and σ_X (σ_{Xr}) denote the average and variance of X (Xr), respectively. $\sigma_{X,Xr}$ is the covariance of X and Xr. c_1 and c_2 are required to stabilize the division.

For our problems, outcomes are the sensing locations provided for each method. Then, we don't have any reconstruction in order to compare different approaches. Thus, an alternative quality analysis based on the variability of simulations conditioned by the obtained sensing locations is proposed. Then we can estimate the mean and standard deviations from simulations for each implemented approach.

Another analysis is to evaluate the global behavior of simulations in relation with the target field. We propose the SNR indicator between the mean of simulations and the reference image and the mean of the SNR of each simulation and the reference image. Finally we propose to use some reconstruction method from data measured at sensing locations like total variation inpainting.

Chapter 3

Preliminary Results

As part of preliminary work we evaluated inverse problems approaches applied to *Geostatistics subsurface channels* and *Imaging*. We organize the preliminary outcomes in the next sections as theoretical or formulation stages and synthetic experimental stages.

In addition, in the current year we are consolidating the developed approaches writing a manuscript related with *OWP* in *2-D* binary channels of permeability.

3.1 On Optimal Well Placement Formulation

There is an interesting interpretation of the eq. (2.6) (in section 2.5), as a problem of maximum information extraction. In this section we summarize this interpretation and study various suboptimal approaches. Furthermore, we propose a field complexity indicator based on *OWP* and show the behavior of this indicator in some typical distributions.

3.1.1 The Equivalent Maximum Information Decision

Adopting the well-known concept of mutual information in information theory [38], we can define the amount of information that provides a rule $f \in \mathbf{F}_K$ as the reduction of the uncertainty after taking the measurements. More precisely, we define the information of f to resolve X as:

$$I(f) \equiv H(X) - H(X^f|X_f) \tag{3.1}$$

where $H(X)$ is the Shannon entropy of the whole array X (or the *a priori* uncertainty) given by:

$$H(X) = -\sum_{x \in A^N} P_x(X = x) \cdot \log P_X(X = x). \tag{3.2}$$

$I(f)$ is a particular case of the well-known mutual information [38], which implies that $I(f) \geq 0$ and $I(f) \leq H(X)$ [1]. Then our OWP problem from eq. (2.6) can be posted, using eq. (3.1), as the problem of maximizing the information to resolve X with K-measurements:

$$f_K^* = \arg\max_{f \in \mathbf{F}_K} I(f). \tag{3.3}$$

To conclude the formulation, we derive a final interpretation of the decision problem in eq. (3.1) and eq. (3.3) using the chain rule of the entropy [38]. The chain rule tells us that for any $f \in \mathbf{F}_K$ we can decompose the joint entropy $H(X)$ as the sum of the marginal entropy of the sensed pixels $H(X_f)$ and the conditional entropy $H(X^f|X_f)$. Using that in eq. (3.1) and then in eq. 3.3 we have that:

$$f_K^* = \arg\max_{f \in \mathbf{F}_K} H(X_f), \tag{3.4}$$

which is the problem of choosing the K-positions that has the highest *a-priori* joint entropy. Then our original problem of minimizing the posterior Shannon entropy after the selection of the sensing positions in eq. (2.6) is equivalent to the problem of finding the K positions that maximizes the information to resolve X in eq. 3.3, which is also equivalent to find the subset of K-measurements that maximizes the *a-priori* uncertainty before taking the measurements in eq. (3.4).

3.1.2 Resolvability Capacity of $\{X_{\mathrm{i}}\}$

Before moving to next sections, it is interesting to analyze an indicator of the complexity of the field in terms of the capacity to resolve its uncertainty with K oriented measurements. We define the resolvability capacity of $\{X_{\mathrm{i}}\}$ with K-measurements as:

$$\mathcal{C}_K \equiv \frac{I(f_K^*)}{H(X)} \in [0, 1]. \tag{3.5}$$

for all $K \in [N]$. This is the ratio between the information gain of the best sensor-placement rule in eq. (3.3) and the whole entropy of the random field. The two extreme cases are $\mathcal{C}_K = 0$ which implies that the K-measurements produces no reduction in uncertainty, and $\mathcal{C}_K = 1$ which implies that there is no remaining uncertainty after taking the K-measurements, i.e., $H(X^{f_K^*}|X_{f_K^*}) = 0$. In general we can show the following properties:

PROPOSITION 1 For any arbitrary random field $\{X_{\mathrm{i}}\}$:

1. $\mathcal{C}_{K+1} \geq \mathcal{C}_K$, $\forall K \in \{1, \ldots, N-1\}$.

2. $\mathcal{C}_N = 1$ and we can define $\mathcal{C}_0 = 0$.

[1] If $I(f) = H(X)$ implies that $H(X^f|X_f) = 0$, which is equivalent to say that X^f is a deterministic function of X_f [38] and, consequently, the sensing rule f perfectly resolves X with no uncertainty.

Hence, $\{\mathcal{C}_K : K\{1,\ldots,N\}\}$ is a monotonically increasing sequence and its profile gives an insight of how simple or complex is to resolve the information of X in the process of taking K-optimal [2] measurements.

3.1.3 Iterative sub-optimal solution for *OWP*

The *OWP* problem, as presented in Eq.(2.6) is combinatorial and, consequently, impractical for relatively large random images [30]. We propose an iterative solution based on the principle of one-step-ahead sensing. The idea is to construct a sub-optimal sensing rule in an incremental fashion to reduce the complexity of the decision algorithm (to something polynomial in the size of the problem) and, therefore, implementable [42, 22, 17].

Let us begin with $K = 1$. In this context the *OWP* problem reduces to find one position in the array solution of:

$$(\mathrm{i}_1^*) = \arg\max_{(\mathrm{i})\in[N]} H(X_{\mathrm{i}}) \tag{3.6}$$

by Eq.(2.2).

The idea of the one-step ahead approach implies to fix i_1^* and then find the next position in $[N]\setminus\{\mathrm{i}_1^*\}$ solution of:

$$(\mathrm{i}_2^*) = \arg\max_{\substack{(\mathrm{i})\in[N]\\(\mathrm{i})\neq(\mathrm{i}_1^*)}} H(X_{\mathrm{i}}|X_{\mathrm{i}_1^*}) \tag{3.7}$$

Then the problem in Eq.(3.7) finds the point that maximizes the *a priori* uncertainty (conditioning to previous measured data, in this case (i_1^*)), which is the simplest principle to implement.

Hence, iterating this inductive (sensing) rule using the chain-rule of the entropy [38], we have that the k-measurement (after taking $(\mathrm{i}_1^*), (\mathrm{i}_2^*), \ldots, (\mathrm{i}_{k-1}^*)$ is the solution of:

$$(\mathrm{i}_k^*) = \arg\max_{\substack{(\mathrm{i})\in[N]\\(\mathrm{i})\neq(\mathrm{i}_p^*)\\p=1,\ldots,k-1}} H(X_{\mathrm{i}}|X_{\mathrm{i}_1^*},\ldots,X_{\mathrm{i}_{k-1}^*}) \tag{3.8}$$

Therefore with this sequence of ordered data points $\{(\mathrm{i}_k^*) : k = 1,\ldots,K\}$, for every $K \in [N]$ we can construct the iterative sensor-placement rule $\tilde{f}_K^* \in \mathbf{F}_K$ by $\tilde{f}_{K(1)}^* = (\mathrm{i}_1^*), \tilde{f}_{K(2)}^* = (\mathrm{i}_2^*), \ldots, \tilde{f}_K^*(k) = (\mathrm{i}_k^*)$.

Again the first problem minimizes the posterior uncertainty after taking the next measurement, the second maximizes the information gain of the next measurement, and the last finds the point that maximizes the *a-priori* uncertainty (conditioning to previous data, in this case $X_{\mathrm{i}_1^*}$), which is the simplest principle to implement.

[2] in the information theoretic meaning.

Concerning the information of this iterative scheme, we can state that:

PROPOSITION 2 The information of $\tilde{f}_K^*$ to resolve the field X (using the definition in eq. 3.1) is given by:

$$\begin{aligned} I(\tilde{f}_K^*) &= H(X) - H(X^{\tilde{f}_K^*}|X_{\tilde{f}_K^*}) \\ &= H(X_{\mathrm{i}_1^*}) + H(X_{\mathrm{i}_2^*}|X_{\mathrm{i}_1^*}) + \cdots + H(X_{\mathrm{i}_K^*}|X_{\mathrm{i}_1^*}, \ldots, X_{\mathrm{i}_{k-1}^*}). \\ &= H(X_{\mathrm{i}_1^*}, \ldots, X_{\mathrm{i}_{k-1}^*}, X_{\mathrm{i}_K^*}), \end{aligned} \tag{3.9}$$

and, consequently, the information gain of this iterative scheme is *additive* in the sense that:

$$I(\tilde{f}_K^*) - I(\tilde{f}_{K-1}^*) = H(X_{\mathrm{i}_K^*}|X_{\mathrm{i}_1^*}, \ldots, X_{\mathrm{i}_{k-1}^*}) \geq 0. \tag{3.10}$$

Note that the information gain from $K-1$ to K measurements in eq. (3.10) derives directly from the solution of eq. (3.8).

3.1.4 Resolvability Capacity of the Iterated Principle

It is simple to show that the optimal solution f_K is better than the iterative solution $\tilde{f}_K^*$ in the sense of information to resolve X. More precisely, for all $K \in \{1, \ldots, N\}$

$$I(\tilde{f}_K^*) \leq I(f_K^*). \tag{3.11}$$

Consequently, a concrete way to evaluate how much we loss in the reduction of the uncertainty between the combinatorial optimal scheme $\{f_K : K \geq 1\}$ and the practical iterative scheme $\left\{\tilde{f}_K^* : K \geq 1\right\}$, is given by the difference between $\mathcal{C}_K$ in eq. 3.5 and

$$\tilde{\mathcal{C}}_K \equiv \frac{I(\tilde{f}_K^*)}{H(X)} \in [0, 1]. \tag{3.12}$$

We conjecture that the information loss $(\mathcal{C}_K - \tilde{\mathcal{C}}_K)_{K=1,\ldots,N}$ is proportional to how much spatial dependency is presented in the joint distribution of field X. In one extreme, it is simple to prove that for a field with no inter-pixel (spatial) dependency, i.e., $P_X(X)_{(\mathrm{i})\in[N]}) = \Pi_{(\mathrm{i})\in[N]} P_{X_\mathrm{i}}(X_\mathrm{i})$, we have that:

$$\mathcal{C}_K = \tilde{\mathcal{C}}_K, \tag{3.13}$$

and furthermore, the iterative solution is optimal: $\tilde{f}_K^* = f_K^*$ for all K.

3.2 On Regionalized variables with spatial dependence

At this point we worked either without any kind of spatial dependence for random variables or only with very simple assumptions over it dependence for regionalized random variables. Currently we have interest in to develop and complement a formalization and practical reduction for it kind of dependence.

3.2.1 *MRF* models and *Cliqué* structure estimation

When the spatial dependence is headed by a markovian property we can reduce the conditional probabilities by only considering conditionals on the *Cliqué* of the variable of interest:

$$P(X_{\mathrm{i}}|X^{\mathrm{i}}) \equiv P(X_{\mathrm{i}}|X_{CL(\mathrm{i})}) \tag{3.14}$$

Where $CL(\mathrm{i})$ correspond to the *Cliqué* associated to the random variable X_{i} at the position i.

3.2.2 *Cliqué* estimation and Mutual Information

Given again a random field $X = \{X_{\mathrm{i}} : \mathrm{i} \in [N]$, our current goal is to define the multi-point conditional probability in terms of two-point conditional probability:

$$P(X_{\mathrm{i}}|X^{\mathrm{i}}) = f(\{P(X_{\mathrm{i}}|X_k);\ \forall k \in [N] \setminus \mathrm{i}\}) \tag{3.15}$$

Using Bayes rule we can write the following expression:

$$\begin{aligned}
P(X_{\mathrm{i}}|X^{\mathrm{i}}) = \frac{P(X_{\mathrm{i}}, X^{\mathrm{i}})}{P(X^{\mathrm{i}})} = \frac{P(X)}{P(X^{\mathrm{i}})} &= \frac{P(X^{\{\mathrm{i},k\}}|X_{\{\mathrm{i},k\}})P(X_{\{\mathrm{i},k\}})}{P(X^{\{\mathrm{i},k\}}|X_k)P(X_k)} \\
&= \frac{P(X^{\{\mathrm{i},k\}}|X_{\{\mathrm{i},k\}})}{P(X^{\{\mathrm{i},k\}}|X_k)}P(X_{\mathrm{i}}|X_k)
\end{aligned} \tag{3.16}$$

Eq. (3.16) shows that $P(X_{\mathrm{i}}|X^{\mathrm{i}})$ can be rewritten in terms of $P(X_{\mathrm{i}}|X_k)$ for any k in the complement subset. We can use this property to derive a combination of $P(X_{\mathrm{i}}|X_k)$ that is equal to multi-point conditional probability.

$$P(X_{\mathrm{i}}|X^{\mathrm{i}})^{|I_{X^{\mathrm{i}}}|} = \prod_{k \in I_{X^{\mathrm{i}}}} \frac{P(X^{\{\mathrm{i},k\}}|X_{\{\mathrm{i},k\}})}{P(X^{\{\mathrm{i},k\}}|X_k)}P(X_{\mathrm{i}}|X_k) \tag{3.17}$$

Where $I_{X^{\text{i}}}$ represent the set of indexes of X^{i}. Using the $|I_{X^{\text{i}}}|$th-root in both sides of eq. (3.17), we have:

$$\underbrace{P(X_{\text{i}}|X^{\text{i}})}_{\text{multi-point prob.}} = \prod_{k\in I_{X^{\text{i}}}} \underbrace{\left(\frac{P(X^{\{\text{i},k\}}|X_{\{\text{i},k\}})}{P(X^{\{\text{i},k\}}|X_k)}\right)^{\frac{1}{|I_{X^{\text{i}}}|}}}_{\alpha_{\{\text{i},k\}}} \underbrace{P(X_{\text{i}}|X_k)}_{\text{two-point prob.}}{}^{\frac{1}{|I_{X^{\text{i}}}|}} \tag{3.18}$$

That is what we wanted to get, as it was stated in eq. (3.15).

3.2.3 Multi-point and Two-Point Mutual Information relation

Applying the definition for mutual information and Eq. (3.18) we can write the following relations:

$$\begin{aligned}
&I(X_{\text{i}};X^{\text{i}})\\
&= H(X_{\text{i}}) - H(X_{\text{i}}|X^{\text{i}})\\
&= H(X_{\text{i}}) + \sum_{\{X_{\text{i}},X^{\text{i}}\}\in A^{|X|}} P(X_{\text{i}},X^{\text{i}}) \log P(X_{\text{i}}|X^{\text{i}})\\
&= H(X_{\text{i}})\\
&\quad + \frac{1}{|I_{X^{\text{i}}}|}\sum_{k\in I_{X^{\text{i}}}} \sum_{\{X_{\text{i}},X^{\text{i}}\}\in A^{|X|}} P(X_{\text{i}},X^{\text{i}}) \big(\log P(X^{\{\text{i},k\}}|X_{\{\text{i},k\}})\\
&\qquad\qquad + \log P(X_{\text{i}}|X_k)\\
&\qquad\qquad - \log P(X^{\{\text{i},k\}}|X_k)\big)\\
&= H(X_{\text{i}}) + \frac{1}{|I_{X^{\text{i}}}|}\sum_{k\in I_{X^{\text{i}}}} \Bigg\{ \sum_{\{X_{\text{i}},X^{\text{i}}\}\in A^{|X|}} P(X_{\text{i}},X^{\text{i}}) \log P(X^{\{\text{i},k\}}|X_{\{\text{i},k\}})\\
&\qquad\qquad + \sum_{\{X_{\text{i}},X^{\text{i}}\}\in A^{|X|}} P(X_{\text{i}},X^{\text{i}}) \log P(X_{\text{i}}|X_k)\\
&\qquad\qquad - \sum_{\{X_{\text{i}},X^{\text{i}}\}\in A^{|X|}} P(X_{\text{i}},X^{\text{i}}) \log P(X^{\{\text{i},k\}}|X_k)\Bigg\}\\
&= H(X_{\text{i}}) - \frac{1}{|X^{\text{i}}|}\sum_{k\in I_{X^{\text{i}}}} \big(H(X_{\text{i}}|X_k) + H(X^{\{\text{i},k\}}|X_{\{\text{i},k\}}) - H(X^{\{\text{i},k\}}|X_k)\big)
\end{aligned} \tag{3.19}$$

Note that $H(X_{\text{i}}|X_k) + H(X_{\{\text{i},k\}}|X^{\{\text{i},k\}}) - H(X_k|X^{\{\text{i},k\}})$ can be rewritten using mutual information property, $H(X|Y) = H(X) - I(X;Y)$.

$$\begin{aligned} &H(X_\mathrm{i}|X_k + H(X^{\{\mathrm{i},k\}}|X_{\{\mathrm{i},k\}}) - H(X^{\{\mathrm{i},k\}}|X_k) \\ &\quad = H(X_\mathrm{i}) - I(X_\mathrm{i};X_k) + H(X^{\{\mathrm{i},k\}}) - I(X_{\{\mathrm{i},k\}};X^{\{\mathrm{i},k\}}) - H(X^{\{\mathrm{i},k\}}) + I(X_k;X^{\{\mathrm{i},k\}}) \\ &\qquad = H(X_\mathrm{i}) - I(X_\mathrm{i};X_k) - I(X_{\{\mathrm{i},k\}};X^{\{\mathrm{i},k\}}) + I(X_k;X^{\{\mathrm{i},k\}}) \end{aligned} \tag{3.20}$$

Then, we can sort the expression of Eq. (3.19) in three kinds of components:

$$\begin{aligned} I(X_\mathrm{i};X^\mathrm{i}) =& H(X_\mathrm{i}) \\ & - \frac{1}{|X^\mathrm{i}|}\sum_{k\in I_{X^\mathrm{i}}} H(X_\mathrm{i}) - I(X_\mathrm{i};X_k) - I(X_{\{\mathrm{i},k\}};X^{\{\mathrm{i},k\}}) + I(X_k;X^{\{\mathrm{i},k\}}) \\ =& \underbrace{\left[\frac{1}{|X^\mathrm{i}|}\sum_{k\in I_{X^\mathrm{i}}} I(X_\mathrm{i};X_k)\right]}_{\text{one and two-point statistics}} \\ & + \underbrace{\left[\frac{1}{|X^\mathrm{i}|}\sum_{k\in I_{X^\mathrm{i}}} I(X_{\{\mathrm{i},k\}};X^{\{\mathrm{i},k\}})\right]}_{\text{higher order statistics}} - \underbrace{\left[\frac{1}{|X^\mathrm{i}|}\sum_{k\in I_{X^\mathrm{i}}} I(X_k;X^{\{\mathrm{i},k\}})\right]}_{\text{redistributable term}} \end{aligned} \tag{3.21}$$

3.3 On Optimal Well Placement Performance

Based on our preliminary theoretical formalization for the near optimal measurements location, in the characterization of subsurface channelized fields of permeability, we implemented an iterative *OWP* solution of eq. (3.9) and provided a comparison between *OWP* and structured/randomized sampling schemes.

3.3.1 Implementation of *OWP* to *2-D* binary channels

Our scheme consider the stages exposed in the fig. 3.1

3.3.2 Data Base

The Database used in our preliminary experiments considered the next constrains:

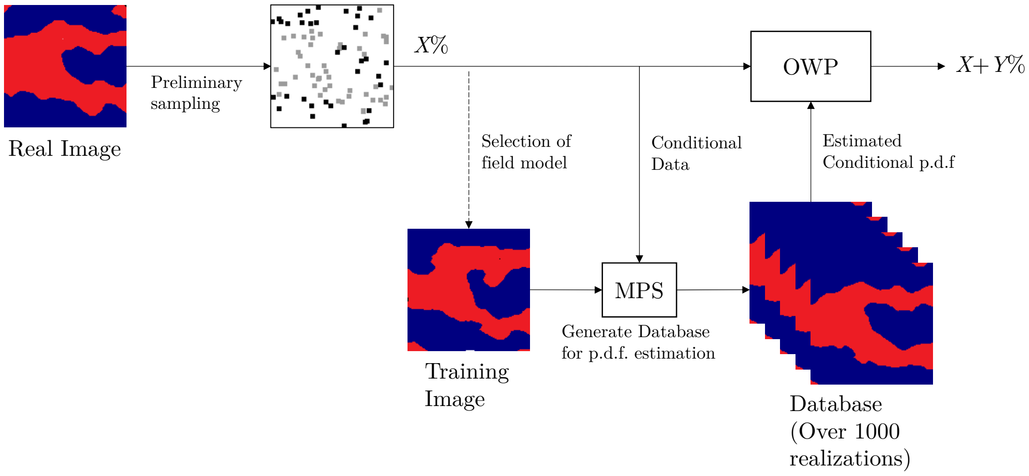


Figure 3.1: Proposed Inference system.

- Nature → Discrete random field
- Dimensionality → *2-D*
- Variable → Permeability
- Dimensions → 200×200 px
- Type → Bi-categorical
- Source → synthetic realizations

3.3.3 Experiments

We considered the next approach to evaluate the sampling design using the proposed database:

- Iterative *OWP* version.
- Combined sampling scheme (uniform sampling (structured) + *OWP*) with an initial presampled positions.
- 200 *MPS* realizations for statistics estimation.

For example, from database we select the next image as the realization image describing the actual channels structure in the subsurface:

Figure 3.2: Realization Image.

First, we used an initial uniform structured sampling [3]and then we applied OWP approach. We overall sampling consisted of the 2% samples from the available positions in the field (realization image). Our modified OWP sampling take into account the next heuristics rules:

- Selection of closest conditionals for **each available position for sampling** (9 nearest point in preliminary heuristic implementation).
- Estimation of conditional probabilities by histograms for *MPS* simluations.
- Estimation of entropy maps.
- Selection of maximal entropy candidates and heuristic corrections.

3.3.4 Experimental Results for OWP

Our experimental setting considered binary images that represent permeability fields. The experiments consisted on the comparison between an equally spaced sensing system and a mixed approach that includes equally spaced measurements and measurements given by the iterative algorithms proposed in eq. (3.9). The objective of this comparison was to contrast the effect that sensing design introduces on the *posteriori* entropy map of the field X. For that, simulations *MPS* tool is also evaluated (by SNESIM algorithm)[36]. The *MPS* technique utilizes a training image, provided by expert knowledge, to perform simulated images that conserve the patterns present in the reference image. In addition, the *MPS* method can be conditioned by measured data.

The first approach consisted on an equally spaced samples of the actual realization field. The second one is described in Fig. 3.1. It required a small portion of preliminary data given by the first method and the other portion sequentially obtained by the application of Eq.(3.8). Conditional probability *PDF*s were estimated using the representative database generated from the simulation algorithm [36, 24], which uses *MPS*. Using a frequentist approach, the probability that current pixel takes a specific value is estimated from the simulated data, for each value of the alphabet $\mathcal{A}$ (in this case, $|\mathcal{A}| = 2$). Doing this operation for each position within f^c, we obtained an estimation of:

$$\mathbb{P}(X_{\{\mathrm{i}\}}|X_f), \quad \forall \mathrm{i} \in f^c \tag{3.22}$$

Here, we computed the conditional entropy, required to apply the OWP algorithm proposed in Eq.(3.8). The analysis of the results relies on the behavior that the *posteriori* entropy takes as we observe in Fig. 3.3. In this case, white pixels represents a high uncertainty zones, while black pixels means a complete knowledge of the value that the pixel is taking. As expected, for the scheme in which OWP is included, the uncertainty is significantly reduced with respect to the other method. Both methods use the same amount of measurements (2% of all $N = 200 \cdot 200$ pixels), but in the one presented in

[3]Needed to the right representation of bad conditioned zones, required for the use of *MPS* as a posterior inference system

this proposal, we used $X\%$ uniformly sampled and $Y\%$ by our *OWP* method, where $X+Y=2\%$.

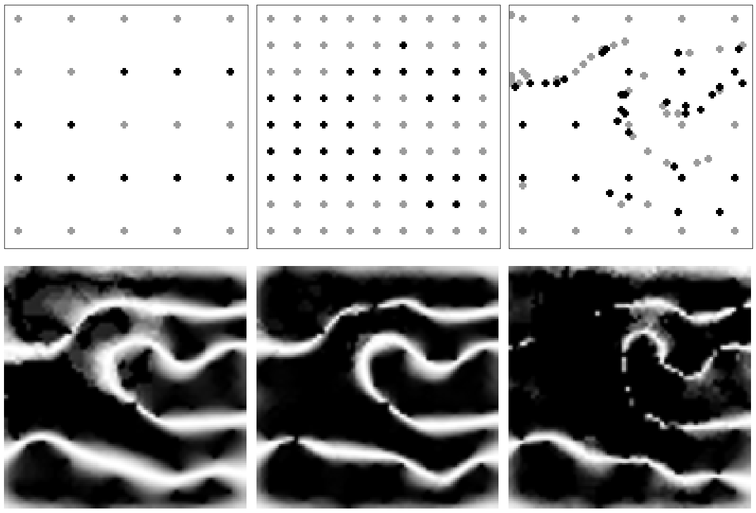


Figure 3.3: Preliminary experiment for adaptive sampling design.
Both measurement systems are shown with their *posteriori* entropy map respectively. In the left column partial measurement process (25 uniformly sampled measurements) is shown previous to the *OWP* application. In the middle one, equally spaced sampling method is presented for the 2% of samples. Finally, in the right column, the result after applying *OWP* measurements is shown.

Next step was to generate simulations of the media using *MPS* tool and using the samples provided by the various methods explained above. Using the samples acquired by different sampling schemes, we explored if there is a regime that combines equally spaced data sampling with data sampled through our method, that is near optimal from the point of view of entropy reduction, average reconstruction error, or variability of the *MPS* simulations. Two metrics were proposed to measure performance: Signal-to-Noise ratio (SNR) between simulations and the reference image in order to analyze the quality of the realizations; and standard deviation of simulations, to measure the uncertainty introduced.

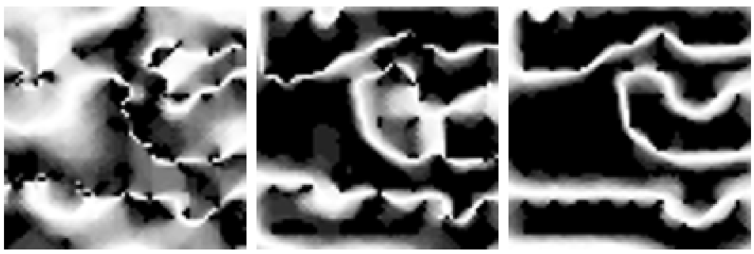


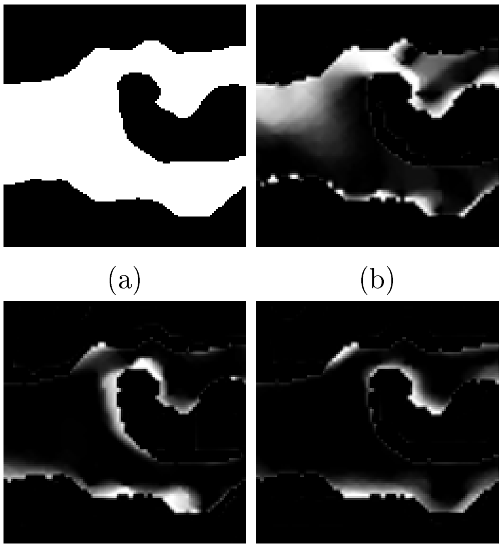
Figure 3.4: Standard deviation maps for different sampling schemes.
From left to right we have: full OWP, 60% equally spaced sampling and 40% OWP (optimal from SNR point of view) and a full equally spaced sampling scenario.

We performed 200 simulations for each sampling scheme. As indicated before, we considered the scenario where 2% of samples are available to do the reconstructions based on simulations. The mixed scheme considers $0\%, 20\%, 40\%, 60\%, 80\%$ and 100% of equally

Table 3.1: Summary of performance for the *OWP* experiment.

$|\sigma|$ corresponds to the total energy present in the standard deviation map.
<SNR> denotes mean *S.N.R* between simulations and *RI*.

Metric	0%	20%	40%	60%	80%	100%
<SNR>	13.9218	20.0376	23.2711	28.1413	22.9547	28.1044
$\|\sigma\|$	18.2311	17.8426	13.0801	11.1337	10.3974	11.605



(a) (b) (c) (d)

Figure 3.5: Mean error for sampling schemes.
a) Reference Image. b) Mean absolute error of full OWP scheme. c) Mean absolute error of 60% equally spaced and 40% OWP scheme. d) Mean absolute error of full equally spaced scheme.

spaced data, and the results are shown in Table 3.1. The percentages shown in the first row make reference to the portion of equally spaced data used by the mixed method. From a <SNR> point of view, the optimal case is obtained when combining 60% of equally spaced data and 40% acquired through OWP method (see Fig. 3.5). For the standard deviation case, optimum is placed in the range of the $80-20\%$ combination (see Fig. 3.4). These results indicated that both kind of data placement methods are complementary and that there exists a trade-off between characterization and uncertainty reduction. Provided results were coherent with the proposed algorithm nature that attempt to reduce global uncertainty on the overall process.

3.4 On Noisy Compressive Sensing Performance

NCS theory has motivated us to study the spatial co-dependencies of the regionalized variable of interest. heretofore, we have investigated in how *MPS* can help in achieving this task and we have initiated an analysis of how the incorporation of spatial dependence improves sparse promoting algorithms.

3.4.1 Statistical Analysis from *MPS*

Using the data base of multichannel MC_1 model we obtained 200 *MPS* realizations from several hard data measurements levels. For each hard data regime, then we estimated second order statistics. As shown in fig. 3.6 the field variance in low rate sampling regimes is higher than in regimes with more measurements. In the extreme case of 0.1% measurements, the variance at most field positions was extremely uncertain. As the hard data measurements increased a reduction of non local variance was observed, thereby the powerful of spatial conditioning from *MPS* was validated.

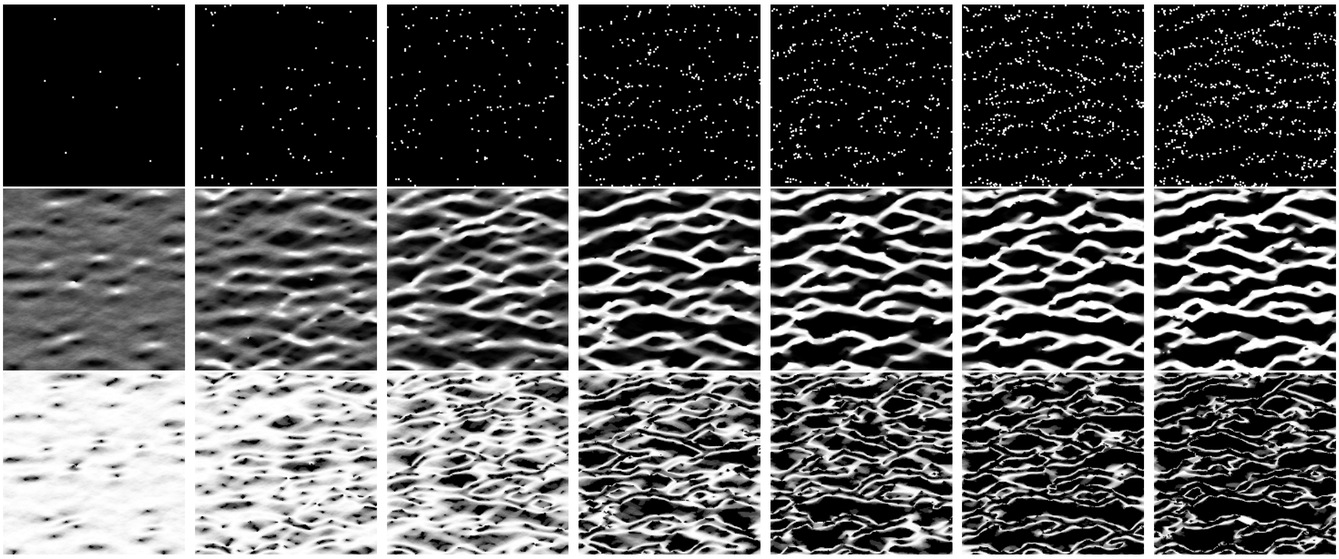


Figure 3.6: Statistics from simulations.
For scenarios considering 0.1 %, 0.5 %, 1 %, 2 %, 3 %, 4 %, and 5 % of hard data measures.

As expected for *MPS*, these preliminary outcomes shown that statistical analysis from *MPS* realizations provides some kind of information about the variability of the value of a non measured pixel from the knowledge of positions and values of hard data measurements. This let us to propose the use of *MPS* realizations as an estimation of the spatial covariance between regionalized variables conforming the stochastic field.

In next subsections we explored several attempts of incorporating information from *MPS* realizations in the *NCS* problem. The main idea was to provide an estimation of covariance matrix and to use it in whitening process.

3.4.2 *NCS*. Naive Approach I. what Covariance???

The most simple assumption was considered that there was no correlation between regionalized variables, and that all variables was modeled incorporating only withe noise with zero mean and the same variance for all the regionalized variables. Under this assumption the variance noise at each individual variable was estimated directly by pixel variance estimation from *MPS* realizations.

At standard *CS* approaches only hard data measurements are considered as input to the reconstruction methods and no one knowledge is given for unmeasured regionalized variables. The incorporation of *MPS* allows no only statistical analysis for global stochastic field, but also the opportunity of considering simulated regionalized variables as noisy virtual measures.

Here most naive *NCS* approach correspond to consider the model from eq. (2.15) but with $C_v^{\frac{-1}{2}}$ equal to the identity matrix of proper size. Thus, using a mixture of hard data measures and simulated *virtual* measures in standard *CS* does not take advantage of the statistical information provided for the proposed *MPS* analysis because all measurements are considered without uncertainty. The above entails that the *CS* reconstruction would be more close to the specific *MPS* realization than to the real target image.

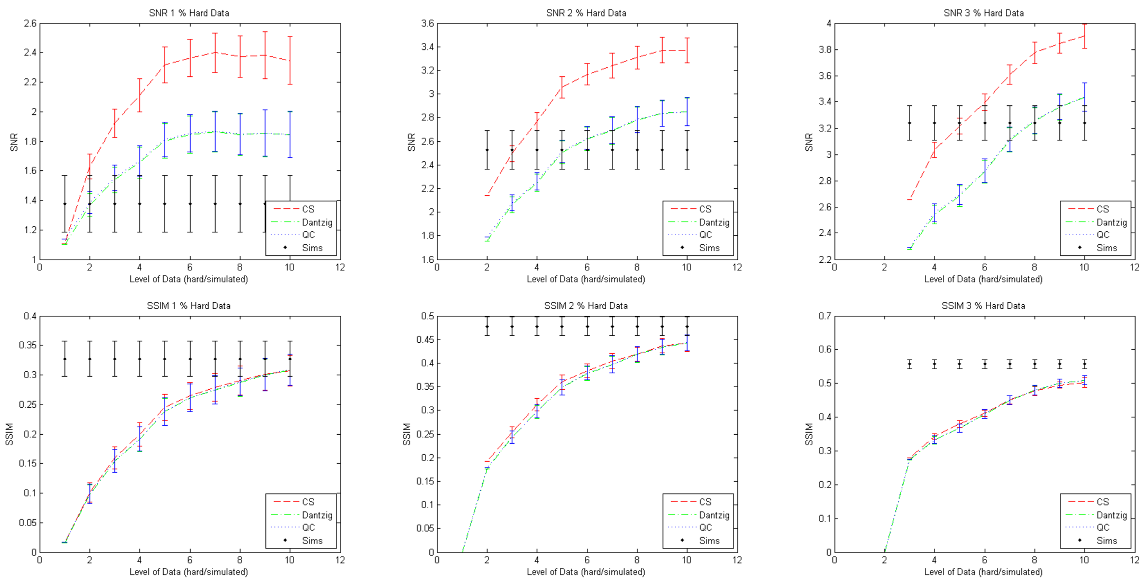


Figure 3.7: Performance Analysis of Naive Approach of *NCS* by only relaxing restrictions on soft data.

From an initial amount of hard data measurements some level of soft virtual measurements are added from simulations and then *CS* and naive *NCS* approaches are applied. Upper row *SNR* analysis, lower row *SSIM* analysis.From left to right : initial 1% of hard data , initial 2% of hard data, and initial 3% of hard data.

As shown in fig. 3.7, partial results on the most naive *NCS* approach present a lower performance than classical *CS*. As reference the performance of *MPS* is shown by the statistics of simulations (mean and variance), *CS* curves consider hard data and soft data as fixed measures while *NCS* allows some uncertainty for these measures. For *SSIM* indicator all reconstructions curves has very close performance, while for *SNR* indicator the classical *CS* achieve a better global performance.

Without incorporating spatial correlations on the regionalized variables we are not taken advantages from *NCS* theory. Including uncertainty on measures only increase the searching space for the optimization algorithms. Thus, our next motivation was including additional constraints promoting desired features on the regionalized variables.

Our next step has corresponded to incorporate proposed amendments from section . Using alternate recovery models promoting sparsity on signal gradients instead on the signal itself.

In figure 3.8 several configurations of hard and virtual data are presented but considering reconstruction approaches using total variation. In terms of visual inspection is possible to appreciate an improvement in the performance of *NCS*. This behavior is confirmed by analysis of the metrics *SNR* and *SSIM* as shown in fig. 3.9.

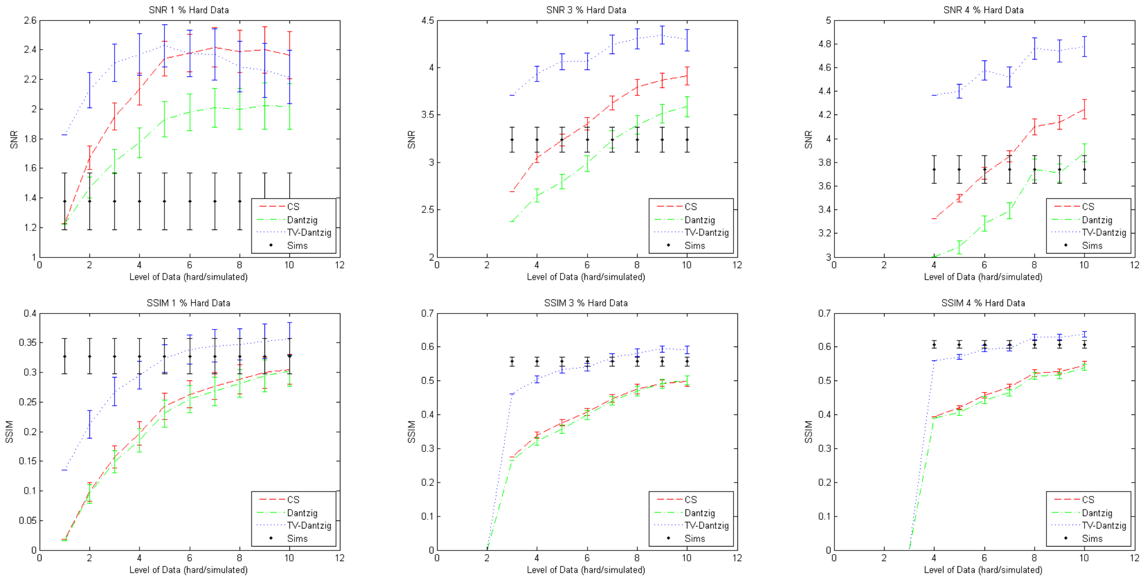


Figure 3.9: Performance Analysis of Naive Approach of *NCS* by only relaxing restrictions on soft data adding *TV* as sparsity promoting approach.

From an initial amount of hard data measurements some level of soft virtual measurements are added from simulations and then *CS* and naive *NCS* approaches are applied. Upper row *SNR* analysis, lower row *SSIM* analysis.From left to right : initial 1% of hard data , initial 3% of hard data, and initial 4% of hard data.

3.4.3 *NCS*. Naive Approach II. Spatial Independence

Here, regionalized variables are considered independent but for each pixel the noise variance is estimated from the empirical variance by the values simulated at this pixel. Thus far we have only qualitative results which account for an apparent improvement in the method's ability to capture something of the structure of original image (figs. 3.10, 3.11, 3.12 and 3.13).

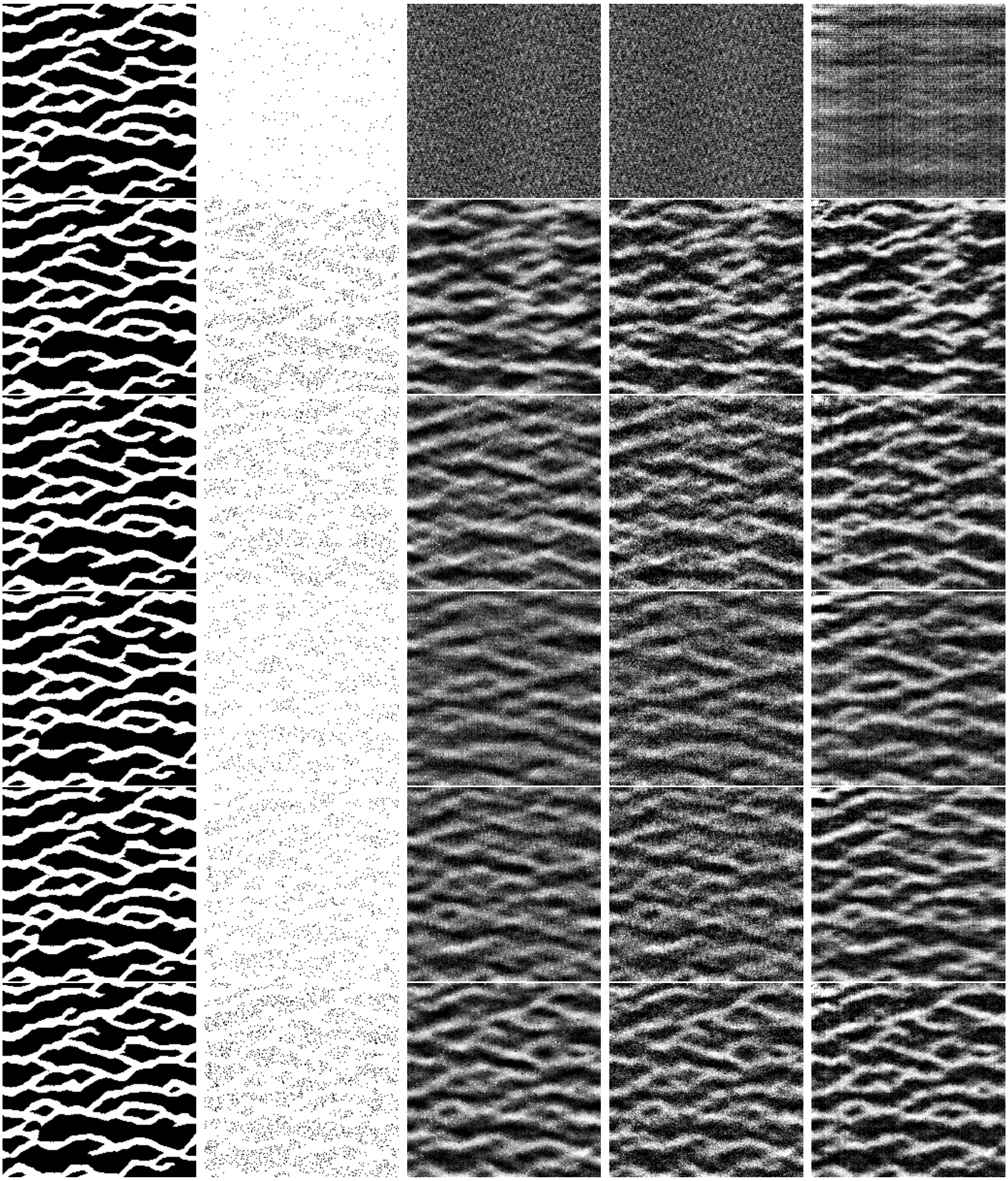


Figure 3.8: Examples of outcomes for *NCS* without considering spatial dependence.
From left to right: Target field image, hard data plus simulated data (HD+SD level), standard *CS* reconstruction for HD+SD, *NCS* for HD+SD by Dantzig selector approach, *NCS* for HD+SD by Dantzig selector and *TV* approach. Rows: Different levels of hard data and simulated data from *MPS* realizations: 1 % *HD* plus 0 % *SD*, 1 % *HD* plus 9 % *SD*, 2 % *HD* plus 8 % *SD*, 4 % *HD* plus 6 % *SD*, 5 % *HD* plus 5 % *SD*, 5 % *HD* plus 10 % *SD*

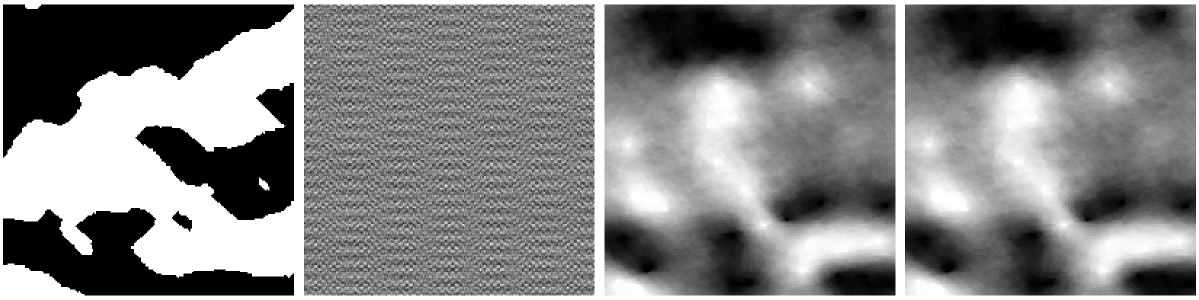


Figure 3.10: Example of *NCS* reconstruction under assumption of independence: Single Channel 1.

Hardata 0.1 %. True Image. Standard *CS*. *NCS* by quadratic constraints. NCS by Dantzig selector.

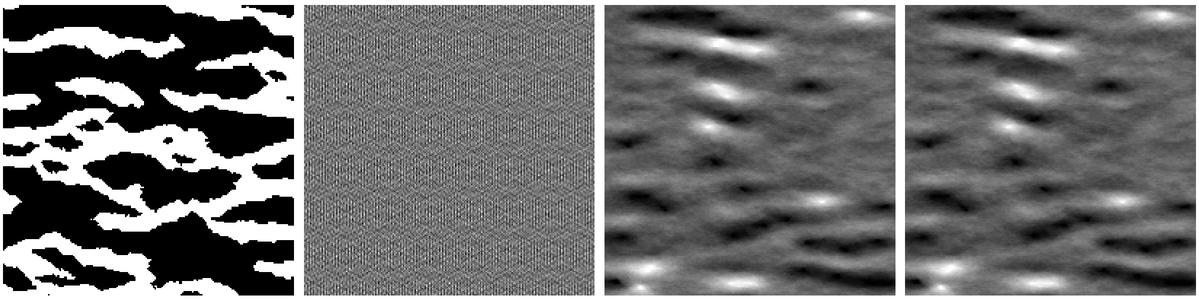


Figure 3.11: Example of *NCS* reconstruction under assumption of independence: Multi Channel 1.

Hardata 0.1 %. True Image. Standard CS. NCS by quadratic constraints. NCS by Dantzig.

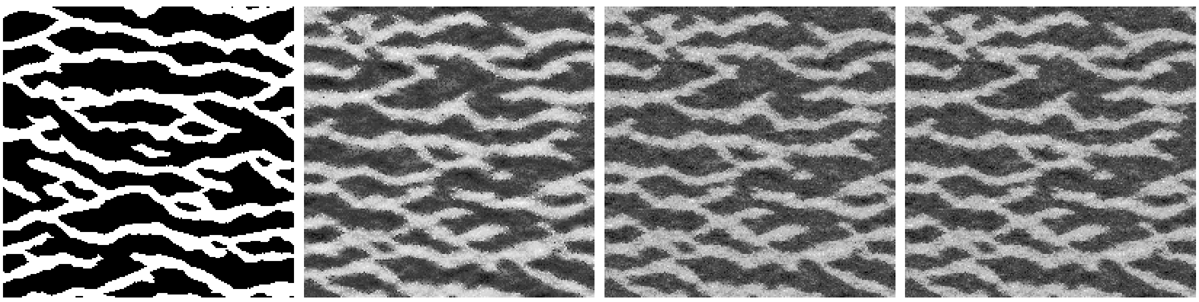


Figure 3.12: Example 1 of *NCS* reconstruction under assumption of independence: Multi Channel 2.

Hardata 0.1 %, Soft data 29.9 %. True Image. Standard CS. NCS by quadratic constraints. NCS by Dantzig selector.

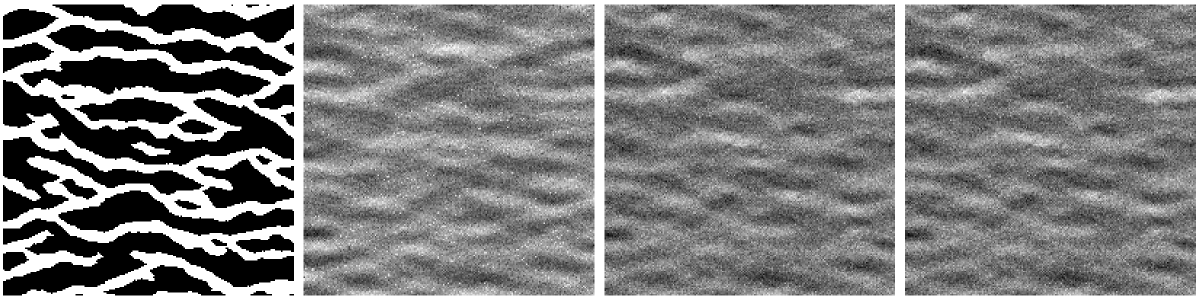


Figure 3.13: Example 2 of *NCS* reconstruction under assumption of independence: Multi Channel 2
Hardata 0.1 %, Soft data 3.9 %. True Image. Standard CS. NCS by quadratic constraints. NCS by Dantzig selector.

3.4.4 *NCS*. Approach III. Full Covariance

Here, regionalized variables are considered with its full spatial dependence. The spatial dependence requires a dense covariance matrix of size N by N demanding more computational resources. Thus far we have only qualitative results which account for an apparent improvement in the method's ability to capture both global structure and deltails of original image from low sampling regimes (figs. 3.14 and 3.15).

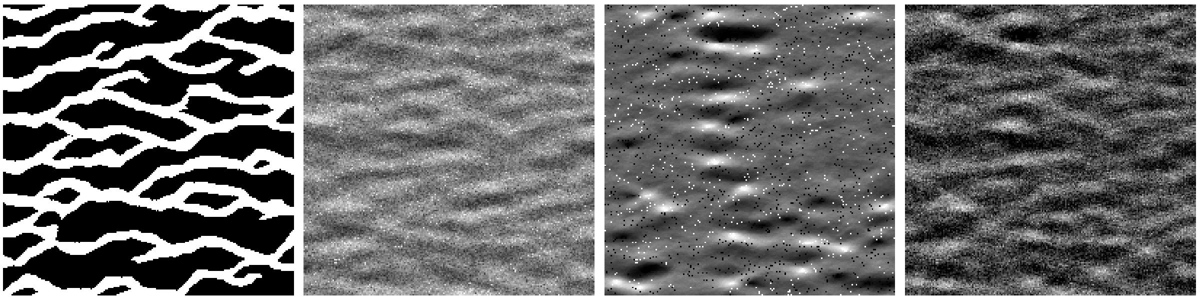


Figure 3.14: Example 1 of *NCS* with full covariance estimation: Multi Channel 2.
Hardata 0.1 %, Soft data 3.9 %. True Image. Standard CS. NCS by quadratic constraints. NCS by Dantzig selector.

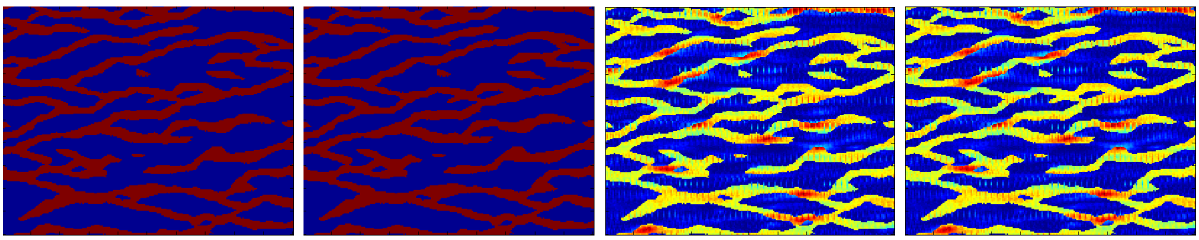


Figure 3.15: Example 2 of *NCS* with full covariance estimation: Multi Channel 2.
Hardata 1 %, Soft data 99 %. True Image. Standard CS. NCS by quadratic constraints. NCS by Dantzig selector.

Chapter 4

Roadmap

In order to achieve the proposed objectives the next stages must be accomplished:

4.1 Proposed stages

4.1.1 Generation of Data Bases

The main topics for this proposal will be studied on a theoretical framework but for evaluation and performance tasks we require a simplified but relevant data set.

Definition of Media

Images denoting *2-D* binary permeability channels will be used as target media. Aforementioned on preliminary sections, this kind of media are common in subsurface characterization and its description provide useful insights into understanding of other subsurface properties. At the outset, we will work under a stationary assumption for the channels.

We will consider 3 kinds of simplified channels: a single channel model and two multichannel models with different level of complexity. These models have the form of *2-D* discrete images representing a realization of a real channel. The realizations of the models are obtained using the standard *SGEMS* program (by *SNESIM* algorithm) and an additional preliminary training image provided by an expert.

Images of reference

From *SNESIM* algorithm we simulate realizations for each kind of channel. Each image have a size of 200 by 200 where each pixel with value 1 represents a high permeability position. Selected images will be used as realizations images (*RIs*) of the actual subsurface channels while others will play the role of training images (*TIs*).

Validation of Database

For the database validation we take into account the preservation of the patterns structure from the reference model and refuse realizations with discontinuities not present in the reference model.

Preliminary Sampling Schemes

In both, *Sensing Placement* analysis and *sparse promoting* methods we require a subset of measurements as preliminary hard data. These preliminary sets work as constraints for *MPS* realizations and we will be obtained by sensing schemes including random, structured and random structured sampling. These schemes will be implemented on *Matlab* framework.

4.1.2 Implementation of parser for MPS

In order to use *MPS* simulations for *NCS* algorithms, we will implement an efficient parser to interact with *SGEMS* software. The proposed parser works on python and interact with *SGEMS* software and *Matlab* codes in a parallel framework.

Multi-Point Simulations

With the database, the preliminary samples, and the developed parser, we will simulate a set of 1000 images for each kind of channel. Each image has a size of 200 by 200.

Statistics Estimation for MPS

For each set of images obtained from *MPS*, we will estimate statistical information about the associated field. In the one side, for *OWP* problem we will use these sets as a source of patterns realizations to approximate joint and conditional *pdfs*. In the other side, for *NCS* these sets will work as a side information source of the spatial correlation for the regionalized variable.

4.1.3 System Modeling for Sparse Promoting Algorithms

The basic considerations required to implement an appropriate *NCS* solver are:

- Whitening Processing: Current *NCS* theory was developed under the assumption of a model with white noise. Therefore, we must apply a whitening pre-processing to adapt the measured signal to the standard implementations of *NCS*.
- Basis Selection: We need to find an appropriate base to promote sparsity or compressibility. Previous work at our research group indicates that *DCT* and *Haar* transforms provide some interesting Basis to evaluate.

Implementation of NCS Algorithms

Based on l_1 *Magic* software, available for *Matlab* platform, we extend and implement *NCS* algorithms oriented to recovery the channelized structure of interest from low acquisition regimes.

- Dantzig Selector Solver: To implement and modify the classical *NCS* method with l_∞ norm regularization, with several definitions and approximations for covariance matrices.
- Basis Pursuit Denoising Solver: To implement and modify the classical *NCS* method with l_2 norm for the error minimization, with several definitions and approximations for covariance matrices.
- Total Variation Solver: To implement and modify approaches based on total variation regularization.

4.1.4 System Modeling for Adaptive Sensing Schemes

To achieve this stage we need to develop the next tasks:

- OWP Formalization: We will analyze the optimal placement problem to the binary channelized structures and develop a theoretical framework to solve it.
- Optimal and Sub-Optimal Approaches: We will study the combinatorial solution associated to the optimal placement and some practical sub-optimal solutions based on iterative approaches.
- Experimental Entropy Estimation: We will estimate statistical information of channelized fields from simulations of *MPS* and training images.

Implementation of OWP Algorithms

Based on the modeling framework developed for *OWP* we will develop the next solvers:

- Entropy Estimation by *MPS*: To implement and modify an empirical frequentist approach for entropy estimation from several *MPS* realizations.
- OWP-MPS Solver: To implement and integrate the statistical approach and the *OWP* solver in a solver to find (near)optimal sensing positions from low sampling acquisition regimes
- OWP-MPSe Solver: To implement and study a framework with re-simulation of *MPS* using previous *OWP* positions as hard data for a sequential *OWP* solver.

4.1.5 Time measurements and Performance Analysis

In order to assess the performance of developed solutions, we will measure time required for the execution of each specific solver and we will study several indicators of overall performance from image reconstruction and entropy reduction.

4.1.6 Writing Preliminary manuscripts

Based on preliminary outcomes we will write the next proposals:

- Sampling selection and *MPS*, an information theoretic approach: We propose the writing of an initial paper of the application of optimal sensing placement theory to the characterization of *2-D* binary channelized structures.
- Qualification Exam Report : We write this thesis proposal supported by previous work and potential contributions of our research topic.

4.1.7 Modification of Proposed Algorithms

From preliminary results and the associated limitations of the classical constrains of *NCS*, we propose the next stages to improve it performance:

- Joint TV-NCS Algorithms: To implement and modify approaches based on total variation regularization and to integrate this solver with *NCS* algorithms
- Variance Matrix Approximation Approach: We will implement approximated versions for the variance matrix obtained from *MPS* because its high dimensionality and complexity impose practical computational constrains. We will modify variance matrix estimation by the incorporation of additional constrains to the spatial correlation in the regionalized variables.

From preliminary results of *OWP*, we propose the next stages to improve its performance:

- Patterns Analysis from TI: We will modify our pattern analysis for *pdfs* estimation bypassing *MPS* by using the training images as a source of pattern occurrences on the channelized structure of interest. We will study several theories developed to address it problem.

- TIPS Formulation: We will implement a training image based pattern search *TIPS*.
- OWP-TIPS Approach: We will modify *OWP* to incorporate *TIPS* approach.

4.1.8 Formalization of Proposed Approaches

To address more general approaches of sampling design and *OWP*, we propose the study of several classical theoretical regionalized fields and propose an indicator of media complexity:

- Resolvability Capacity (*RC*): We will define an indicator of media complexity based on the entropy reduction provided by the incorporation of *OWP* positions to the measured knowledge of the field of interest.
- Gaussian Fields and OWP: We will study several common literature related with optimal sensor placement for continuous regionalized medias.
- Markov Chains Modeling for Binary Maps: We will study some models of spatial dependence based on Markov chains and to implement *Cliqué* based principles in our *OWP* framework.
- *RC* and Medias with Theoretic *pdfs*: We will study the behavior of *RC* under some regionalized fields with classical theoretic *pdfs*.

We will study Adaptive Compressive Sensing (*ACS*) theory and implement some approaches oriented to merge our *OWP* principles and *NCS* based reconstruction solvers in a general framework. As our previous implementations are focused on near optimal estimations we will work on some Near Optimal Adaptive Compressive Sensing (*NoACS*) theoretical results and its possible use in our problem.

Implementation of Near-Optimal Adaptive Compressed Sensing

We will implement versions of *NoACS* focused on *2-D* binary signals by the modification of our preliminary versions of *OWP* and *NCS* solvers.

4.1.9 Writing manuscripts for journals

We will to write a couple of additional publications and to assist some conferences related with our current and future work. In detail, we propose at least the next topics:

- Spatially correlated media characterization, resolvability capacity as a complexity media indicator: We will to write a manuscript about our indicator of media complexity applied to the study of well known regionalized fields.

- Geological Binary Permeability Channels Images Reconstruction by Noisy Compressive Sensing: We will to write a manuscript consolidating our work of *NCS* for channelized structures.
- NoACS for Binary Permeability Channels: We propose the beginning of a manuscript incorporating the ideas developed for *NoACS* applied to regionalized variables.

4.2 Gannt charts

To organize preliminary and future work we provide two separate time tables. The first one oriented to our preliminary work related with this thesis proposal and the second one oriented to the current and future work for the next years.

4.2.1 Preliminary Work

See figure 4.1

4.2.2 Future work

See figure 4.2

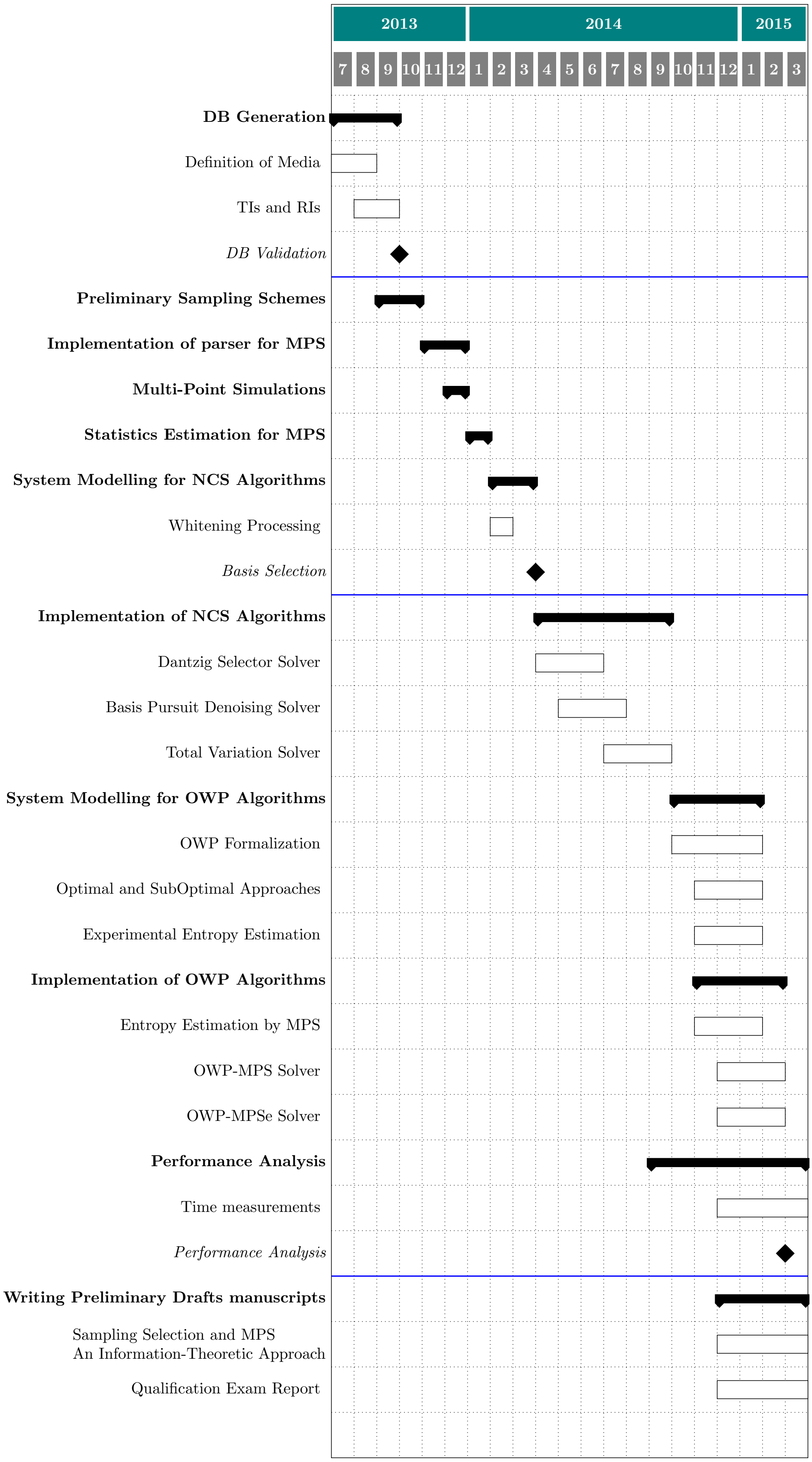


Figure 4.1: Gannt Chart for Preliminary Work

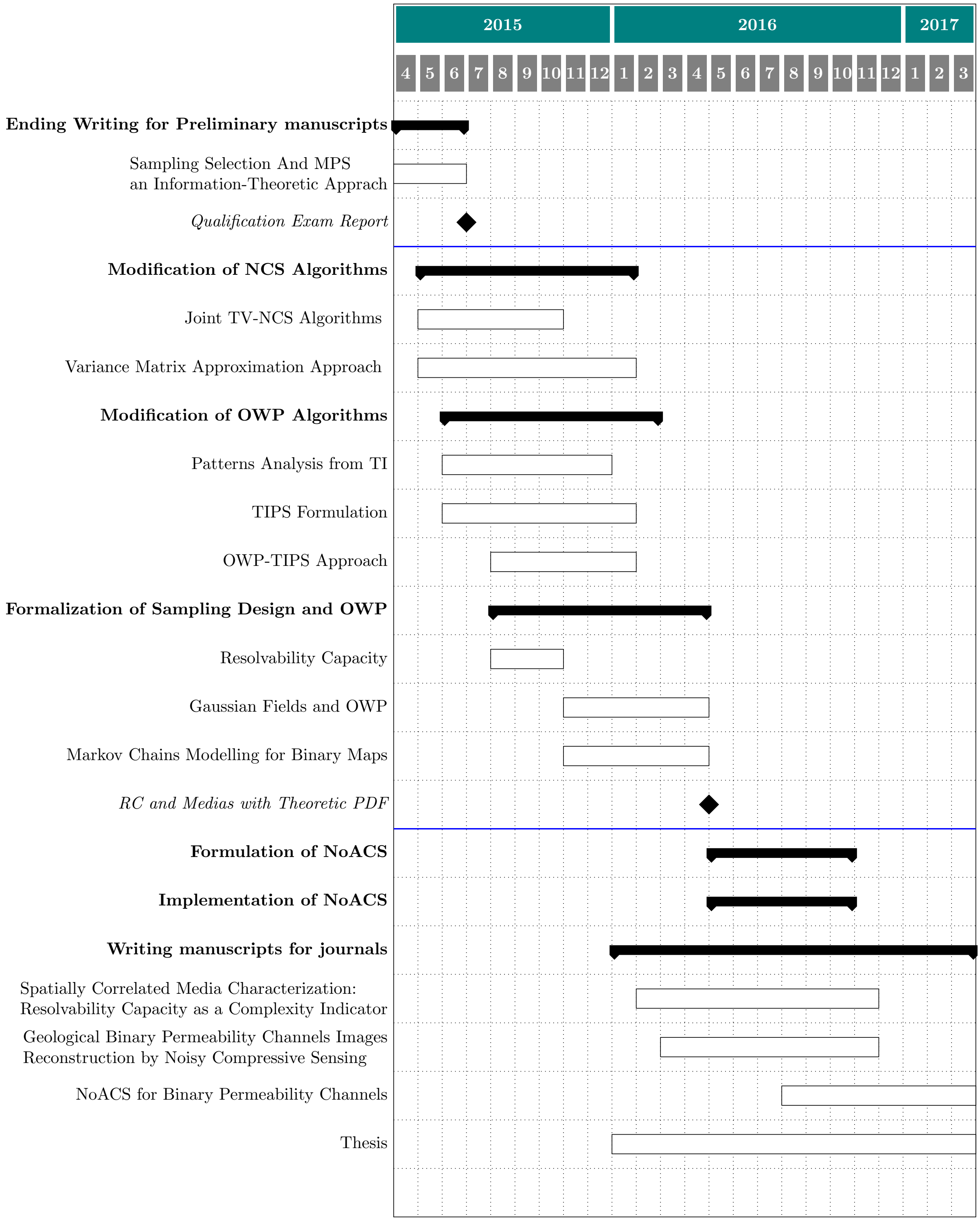


Figure 4.2: Gannt Chart for Current and Future Work

Chapter 5

Expected Outcomes

We have proposed the use of sensing design approaches and sparse promoting regularization methods to solve ill-posed problems of *2-D* binary subsurface channels characterization. We will address this problem by incorporating side information of spatial correlations from training images and *MPS*. We want to extend and to integrate previous outcomes from sparse representation and sensing theory.

Preliminary results give us some idea about the overall impact of proposed approaches. These approaches outperforms the *MPS* outcomes based on *SNESIM* algorithm. Classical *MPS* tools provide only a wide set of plausible realizations of the wanted signal focus on honoring patterns in a training image. However, the proposed methods are focused on recovery actual structure from low acquisition regimes providing a robust mechanism to estimate the field details with less variability.

A key expected contribution rely on the fusion of sensing theory and sparse recovery tools by an adaptive compressive sensing algorithm applied to the characterization of geological facies combined with the integration of geological prior information from the spatial correlation in the regionalized variables under analysis.

On the one side, for *OWP* approaches we propose the *pdf* estimation by the use of pattern occurrence analysis in the training images and from *MPS* realizations. The hypothesis regarding that *MPS* methods generate realizations preserving high order statistics from training images conditioned to hard data allowing an empirical approximation of the statistical information about the field.

On the other side, for *NCS* approaches we propose the use of a set of thousands of realizations of the field by *MPS* to calculate the statistical spatial correlation of the regionalized model. Therefore, we will to estimate covariance matrices. Thus, *MPS* is proposed as an excellent information source to estimate spatial dependence of the actual field conditioned to the *TI* and hard data. Additionally, several definitions of covariance approximations will be analyzed, in order to illustrate the relevance of a full or partial spatial interdependence characterization.

Conclusion

This research proposal focus on: (1) to find the sampling and reconstruction conditions required for the characterization of *2-D* regionalized fields under the assumption of sparsity/compressibility and considering the incorporation of side information related with spatial dependence, (2) to improve previous outcomes in the reconstruction of *2-D* signals for compressible generic ones, and (3) to reduce the variability of the *MPS* realizations providing an useful tool for exploration and characterization tasks. In order to achieve these topics, we propose the use of both sampling design methods and sparse promoting solvers in a novel near optimal adaptive compressive sensing approach. We will adapt, improve, and exploit prior information from training images in channelized structures characterization. Furthermore, we will assess several covariance matrices approximations based on spatial constraints for *NCS*.

It is important to underline that there are not previous work that merge the topics covered in this proposal for the regionalized variables studied. In addition, there are not previous comprehensive analysis developed in this kind of media that exploit spatial constrains in the way proposed. Therefore, from the expected outcomes for this research we could ultimately lead to novel methods of *NoACS* exploiting spatial constrains of binary regionalized fields.

Although the proposed methods will be focused on *2-D* binary channelized structures, the principles and application can be easily extended to other signals with spatial constraints. Depending on the achievement of proposed schedules we would to explore the applicability of our results in others signals at the final stages of this thesis.

Bibliography

[1] J. C. Santamarina and D. Fratta, *Discrete signals and inverse problems : an introduction for engineers and scientists.* Chichester: J. Wiley and sons, 2005. [Online]. Available: http://opac.inria.fr/record=b1105865

[2] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing (3rd Ed.): Principles, Algorithms, and Applications.* Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.

[3] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, 2006.

[4] V. Cherkassky and F. M. Mulier, *Learning from Data: Concepts, Theory, and Methods.* Wiley-IEEE Press, 2007.

[5] Y. C. Eldar and G. Kutyniok, Eds., *Compressed sensing : theory and applications.* Cambridge, New York: Cambridge University Press, 2012. [Online]. Available: http://opac.inria.fr/record=b1134461

[6] P. K. Katanidis, *Introduction to geostatistics: applications to hydrogeology.* Cambridge Univ Press, 1997.

[7] S. Strebelle and T. Zhang, "Non-stationary multiple-point geostatistical models," in *O Leuangthong and C.V. Deutsch (eds.), Geostatistics Banff*, vol. 14, 2004, pp. 235–244. [Online]. Available: http://0-apps.webofknowledge.com.libraries.colorado.edu/full_record.do?product=WOS&search_mode=GeneralSearch&qid=1&SID=4EMumF18jzfjjoowxPB&page=1&doc=5&cacheurlFromRightClick=no

[8] D. S. Oliver, A. C. Reynolds, and N. Liu, *Inverse Theory for Petroleum Reservoir Characterization and History Matching.* Cambridge: Cambridge University Press, 2008. [Online]. Available: http://books.google.com/books?hl=en&lr=&id=oPuy3OfbtfIC&oi=fnd&pg=PR11&dq=Inverse+Theory+for+Petroleum+Reservoir+Charaterization+and+History+Matching&ots=9fIWq86eCn&sig=nrnDTbgY36_iqIEemjzFNw9Xu54http://ebooks.cambridge.org/ref/id/CBO9780511535642

[9] C. Scheidt and J. Caers, "Representing spatial uncertainty using distances and kernels," *Mathematical Geosciences*, vol. 41, no. 4, pp. 397–419, 2009.

[10] J. Onwunalu and L. Durlofsky, "Application of a particle swarm optimization algorithm for determining optimum well location and type," *Computational Geosciences*, vol. 14, no. 1, pp. 183–198, 2009. [Online]. Available: http://dx.doi.org/10.1007/s10596-009-9142-1

[11] Güyagüler Baris, "Optimization of Well Placement and Assessment of Uncertainty," Doctor of Philosophy Thesis, Stanford University, 2002.

[12] J. F. Wellmann, "Information theory for correlation analysis and estimation of uncertainty reduction in maps and models," *Entropy*, vol. 15, no. 4, pp. 1464–1485, 2013.

[13] W. Bangerth, H. Klie, V. Matossian, M. Parashar, and M. Wheeler, "An autonomic reservoir framework for the stochastic optimization of well placement," *Cluster Computing*, vol. 8, pp. 255–269, 2005.

[14] A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg, "Near-optimal sensor placements: Maximizing information while minimizing communication cost," in *Proc. of Information Processing in Sensor Networks (IPSN)*, 2006.

[15] W. Bangerth, H. Klie, M. F. Wheeler, P. Stoffa, and M. Sen, "On optimization algorithms for the reservoir oil well placement problem," *Comp. Geosciences*, vol. 10, pp. 303–319, 2006.

[16] W. Bangerth, H. Klie, M. F. Wheeler, P. L. Stoffa, and M. K. Sen, "On optimization algorithms for the reservoir oil well placement problem," *COMP. GEOSC*, vol. 10, no. 3, pp. 303–319, 2006.

[17] A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, and C. Faloutsos, "Efficient sensor placement optimization for securing large water distribution networks," *Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 516–526, November 2008.

[18] N. Remy, A. Boucher, and J. Wu, *Applied geostatistics with SGeMS : a user's guide*. Cambridge : Cambridge University Press, 2009, formerly CIP.

[19] N. S.-N. Lam, "Spatial interpolation methods: A review," *The American Cartographer*, vol. 10, no. 2, pp. 129–150, 1983. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1559/152304083783914958

[20] G. Matheron, *The Theory of Regionalized Variables and its Applications*, ser. Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau. Ecole des Mines de Paris, 1971.

[21] R. A. Olea, "Sampling design optimization for spatial functions," *Mathematical Ge-*

ology, vol. 16, no. 4, pp. 369–392, 1984.

[22] A. Krause, "Optimizing sensing: Theory and applications," Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, PA, USA, December 2008, aAI3343461.

[23] A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg, "Robust sensor placements at informative and communication-efficient locations," *ACM Transactions on Sensor Networks (TOSN)*, vol. 7, no. 4, February 2011.

[24] J. M. Ortiz and C. V. Deutsch, "Indicator simulation accounting for multiple-point statistics," *Mathematical Geology*, vol. 36, no. 5, pp. 545–565, 2004. [Online]. Available: http://dx.doi.org/10.1023/B%3AMATG.0000037736.00489.b5

[25] L. Holden, R. Hauge, . Skare, and A. Skorstad, "Modeling of fluvial reservoirs with object models," *Mathematical Geology*, vol. 30, no. 5, pp. 473–496, 1998. [Online]. Available: http://dx.doi.org/10.1023/A%3A1021769526425

[26] J.-P. Chilès and P. Delfiner, *Geostatistics : modeling spatial uncertainty*, ser. Wiley series in probability and statistics. New York: Wiley, 1999, a Wiley-Interscience publication. [Online]. Available: http://opac.inria.fr/record=b1098313

[27] F. Guardiano and R. Srivastava, "Multivariate geostatistics: Beyond bivariate moments," *In Soares, A., (ed) Geostatistics-Troia*, vol. 1, pp. 133–144, 1993.

[28] C. Guestrin, A. Krause, and A. Singh, "Near-optimal sensor placements in Gaussian processes," in *International Conference on Machine Learning (ICML)*, August 2005.

[29] A. Krause, R. Rajagopal, A. Gupta, and C. Guestrin, "Simultaneous placement and scheduling of sensors," in *Proc. ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2009.

[30] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research (JMLR)*, vol. 9, pp. 235–284, February 2008.

[31] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, 1st ed. Springer Publishing Company, Incorporated, 2010.

[32] J.-L. Starck, F. Murtagh, and J. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity.* New York, NY, USA: Cambridge University Press, 2010.

[33] N. Boyko, G. Karamemis, V. Kuzmenko, and S. Uryasev, "Sparse Signal Reconstruction: LASSO and Cardinality Approaches," *ise.ufl.edu*, 2011. [Online]. Available: http://www.ise.ufl.edu/uryasev/files/2012/02/tsparse_2_1_Stan2_Gulver_2.pdf

[34] E. Candes and T. Tao, "The Dantzig selector: Statistical estimation when p is

much larger than n," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, Dec. 2007. [Online]. Available: http://www-stat.stanford.edu/~{}candes/papers/DantzigSelector.pdf

[35] A. Chambolle and V. Caselles, "An introduction to total variation for image analysis," *Theoretical Foundations and Numerical Methods for Sparse Recovery*, vol. 9, pp. 263 – 340, 2010. [Online]. Available: http://books.google.com/books?hl=en&lr=&id=MK1E-dch80MC&oi=fnd&pg=PA263&dq=An+introduction+to+Total+Variation+for+Image+Analysis&ots=l6tLYs3fsS&sig=XJGcUhggSl6FDDtGbfnfvhifKyw

[36] T. Huang, D.-T. Lu, X. Li, and L. Wang, "Gpu-based snesim implementation for multiple-point statistical simulation," *Comput. Geosci.*, vol. 54, pp. 75–87, Apr. 2013. [Online]. Available: http://dx.doi.org/10.1016/j.cageo.2012.11.022

[37] J. T. Parker, V. Cevher, and P. Schniter, "Compressive sensing under matrix uncertainties: An Approximate Message Passing approach," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, Nov. 2011, pp. 804–808. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6190118http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6190118

[38] T. Cover and J. Thomas, *Elements of Information Theory*, ser. Elements of Information Theory. Wiley, 2006. [Online]. Available: http://books.google.es/books?id=EuhBluW31hsC

[39] H. Calderón, J. F. Silva, J. M. Ortiz, and A. Egaña, "Reconstruction of channelized geological facies based on {RIPless} compressed sensing," *Computers & Geosciences*, vol. 77, pp. 54 – 65, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0098300415000072

[40] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (2nd Edition)*. Pearson, 2002.

[41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 13, no. 4, pp. 600–612, 2004.

[42] A. Krause and C. Guestrin, "Near-optimal observation selection using submodular functions," in *National Conference on Artificial Intelligence (AAAI), Nectar track*, July 2007.

[43] H. Bui, C. La, and M. Do, "A fast tree-based algorithm for compressed sensing with sparse-tree prior," *Signal Processing*, vol. 108, no. Complete, pp. 628–641, 2015.

[44] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of signal processing*. Cambridge: Cambridge University Press, 2014. [Online]. Available: http://opac.inria.fr/record=b1135612

[45] C. Deutsch and A. Journel, "Geostatistical software library and user's guide, 2nd edition," 1998.

[46] G. B. Arpat and J. Caers, "Conditional simulation with patterns," *Mathematical Geology*, vol. 39, no. 2, pp. 177–203, 2007.

[47] C. Daly and J. Caers, "Multi-point geostatistics–an introductory overview," *First Break*, vol. 28, no. 9, 2010.

[48] P. Mariethoz and P. Caers, *Multiple-point Geostatistics: Stochastic Modeling with Training Images*. Wiley, 2014. [Online]. Available: https://books.google.cl/books?id=kQX2oAEACAAJ

[49] P. Burt and E. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communications*, vol. 31, no. 4, 1983.

[50] P. J. Burt, "Fast Filter Transforms for Image Processing," *Computer Graphics and Image Processing*, vol. 16, no. 1, pp. 20–51, 1981. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0146664X81900927

[51] G. R. Cross and A. K. Jain, "Markov random field texture models," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-5, no. 1, pp. 25–39, Jan 1983.