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 between 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 model by:

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

The problem can be formulated in the following way: 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 the set of parameters are 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.

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

$$\mathbf{Y} = g(\mathbf{X}) + \nu. \tag{1.2}$$

A classical definition for the objective function is the one described in eq. (1.3) :

$$G(\mathbf{Y}, \mathbf{X}, g(\cdot)) = \|\mathbf{Y} - g(\mathbf{X})\|_p \,, \tag{1.3}$$

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

1.2 Problem Relevance

1.2.1 Inverse Problems and Sampling

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

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

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 characterize the band limited signals. It is the principle behind the many used 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.

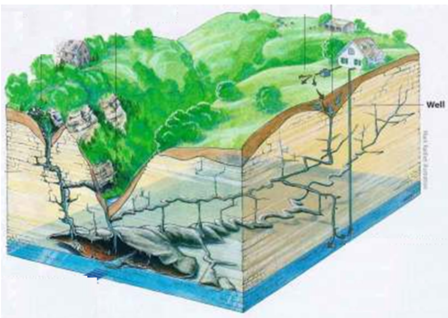

Figure 1.1: Example of a $3D$ geological reservoir.

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 very fine grid in the paths of interest providing better details than seismic data and the uncertainty associated with these measures is very small relative to the seismic sources.

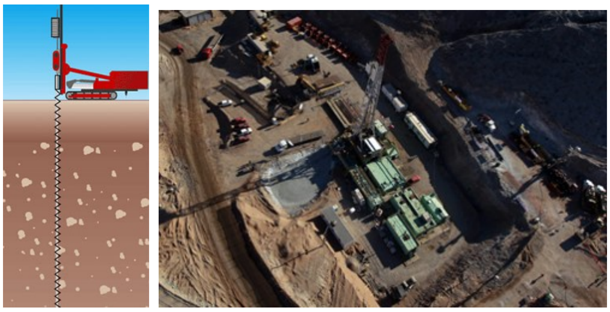

Figure 1.2: Left: Example of a well sampling scheme. Right: Example of an actual well sampling system.

In the case of production data, the information is acquired along the production process in production wells. Production data takes into account global factors related with features of reservoir that makes a great difference with the other sensing modalities. Production history 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 [11, 13, 14, 15, 8, 10]. 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 very 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 also consider and could take advantages of sensing design and inference tools. While blast hole drilling systems has been focused on efficient drilling instead on 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 as can be seen in 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 maps estimation and map making tasks.

Figure 1.3: Example of a $2D$ 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 on 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).

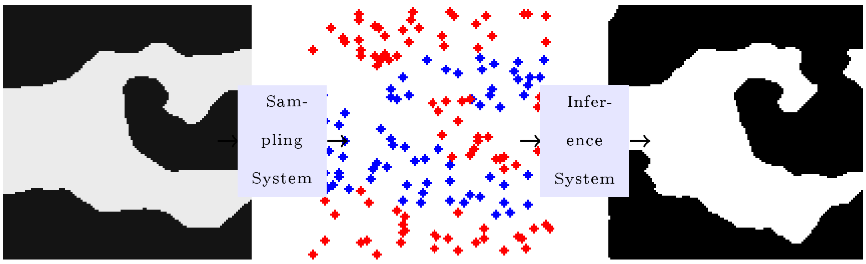

Figure 1.4: Example of a basic scheme of inverse problem related with $2D$ channelized structures characterization.

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 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 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 1.4, Sensing design approaches 1.5 and Recovery/reconstruction methods 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 $2D$ 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 $X = \{X_\mathrm{i} : \mathrm{i} \in \{1, \ldots, N\}\}$.

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].

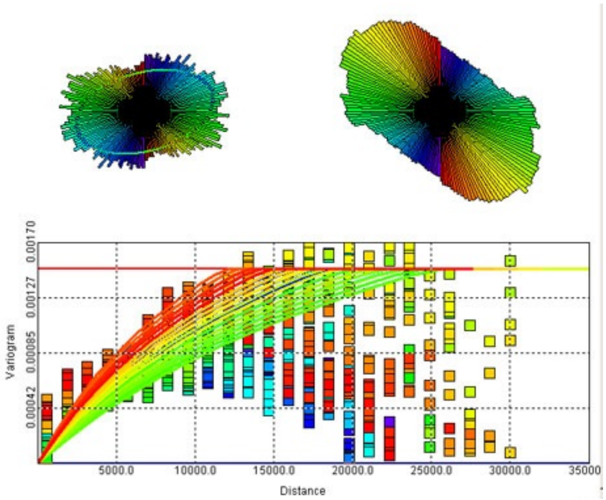

Figure 1.5: Example of a classical variographic analysis.

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 deeper 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

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

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].

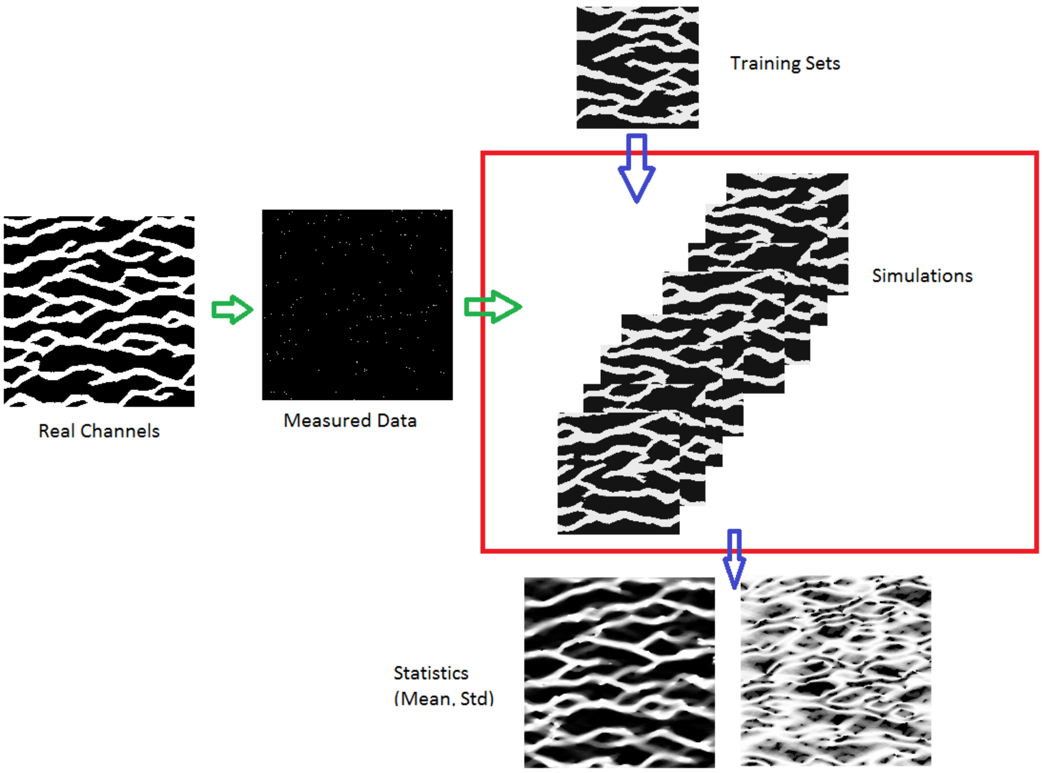

Figure 1.6: Example of a simulation based system.

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. *MPS* 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 reproduces 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 $2D$ or $3D$ valid realization of a field with the same structures than 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. One 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 $X = \{X_{\text{i}} : \text{i} \in \{1, \dots, N\}\}$ 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:

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

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 performance of an inference system or another metric. For example, Fig. 1.7 presents a description of two sensing approaches in a recovery based system.

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

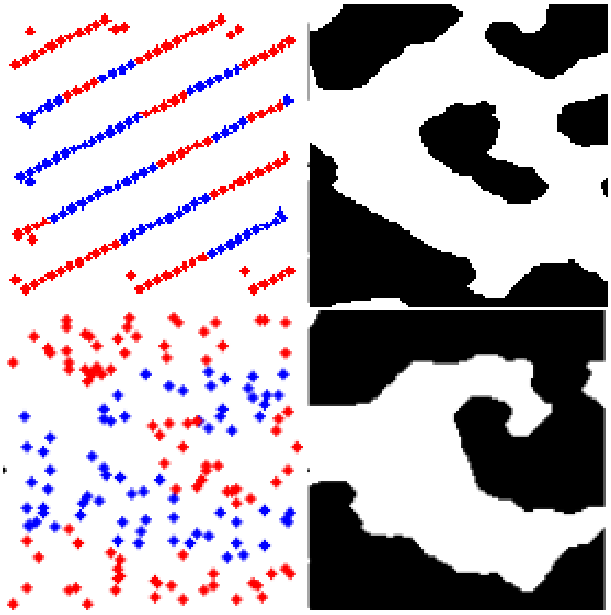

Figure 1.7: Example of sensing design and its effect on reconstruction of the field. Upper row: an arbitrary structured sensing scheme (left) and the achieved reconstruction by the measures at this scheme locations (right). Lower row: a near optimal sensing scheme (left) and the achieved reconstruction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

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

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

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

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

Given an 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 $H(X_{\text{Still No Measured}}|X_{\text{To Measure}})$.

[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

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].

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

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

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

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 $\mathbf{A}$
$\mathbf{A}_\Lambda^\dagger$ denotes the pseudo-inverse of $\mathbf{A}_\Lambda$

OMP	**Thresholding**
initialize: $R = Y_{Obs}$, $\Lambda = \emptyset$ find: $k = \arg\max_j |\langle r, A_j \rangle|$ 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 $|\langle Y_{Obs}, A_j \rangle|$ output: $X = \mathbf{A}_\Lambda^\dagger \cdot Y_{Obs}$

For the noisy case:

$$(P_0 n) \qquad \min_{X \in \mathbb{R}^N} \|X\|_0 \text{ subject to } \|Y_{Obs} - \mathbf{A} \cdot X\|_2 \leq \eta. \tag{1.9}$$

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.

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*:

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

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

Here, ℓ_1-norm is calculated as $\|x\|_1 = \sum |x_i|$.

The main goal of the *CS* theory is to finding 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 $\Lambda \subset \{1,\dots,N\}$ 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 satisfies $Y_{Obs} = \mathbf{A}\cdot X + \xi$ for some $\|\xi\|_2 \le \eta$, 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].

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

In fact, if $\delta_{4S} \le 1/3$ then $C \le 15.41$. In addition, the formulation related with the Eq. (1.14) guaranties 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_0n) 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.

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

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_1n), 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 $2D$ image, an alternate recovery approach is promoting sparsity on gradients of the signal instead on the signal itself. The $2D$ gradient can be calculated from total variations, 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 *TV-BPDN* approaches is defined by:

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

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].