Statistical Fusion of Scientific Images

by

Azadeh Mohebi

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A practical and important class of scientific images are the 2D/3D images obtained from porous materials such as concretes, bone, active carbon, and glass. These materials constitute an important class of heterogeneous media possessing complicated microstructure that is difficult to describe qualitatively. However, they are not totally random and there is a mixture of organization and randomness that makes them difficult to characterize and study. In order to study different properties of porous materials, 2D/3D high resolution samples are required. But obtaining high resolution samples usually requires cutting, polishing and exposure to air, all of which affect the properties of the sample. Moreover, 3D samples obtained by Magnetic Resonance Imaging (MRI) are very low resolution and noisy. Therefore, artificial samples of porous media are required to be generated through a *porous media reconstruction* process. The recent contributions in the reconstruction task are either only based on a prior model, learned from statistical features of real high resolution training data, and generating samples from that model, or based on a prior model and the measurements.

The main objective of this thesis is to some up with a statistical data fusion framework by which different images of porous materials at different resolutions and modalities are combined in order to generate artificial samples of porous media with enhanced resolution. The current super-resolution, multi-resolution and registration methods in image processing fail to provide a general framework for the porous media reconstruction purpose since they are usually based on finding an estimate rather than a typical sample, and also based on having the images from the same scene – the case which is not true for porous media images.

The statistical fusion approach that we propose here is based on a Bayesian framework by which a prior model learned from high resolution samples are combined with a measurement model defined based on the low resolution, coarse-scale information, to come up with a posterior model. We define a measurement model, in the non-hierachical and hierarchical image modeling framework, which describes how the low resolution information is asserted in the posterior model. Then, we propose a posterior sampling approach by which 2D posterior samples of porous media are generated from the posterior model. A more general framework that we propose here is asserting other constraints rather than the measurement in the model and then propose a constrained sampling strategy based on simulated annealing to generate artificial samples.

Acknowledgments

First of all, I would like to express my deep gratitudes to the merciful and almighty God for his everlasting bestowals.

I would like to extend my sincere appreciations to my supervisor, Professor Paul Fieguth, for his instructive role in my thesis. He is not only a benign teacher but also a very knowledgeable researcher. I owe him this thesis for his wise, profound and clever insights in our weekly meetings. He thought me how to analyze a problem thoroughly and precisely and how to determine possible solutions to solve the problem. His critical way of thinking opened a new horizon for me to approach a problem from different perspectives.

I would like to thank my committee members: Professor Modersitzki, Professor Peemoeller, Professor Ioannidis, and Professor Clausi. They have certainly provided me with valuable and insightful comments to promote this thesis. I would like to thank Professor Ioannidis, specifically, for his effective feedback and fruitful remarks on my research.

My deep gratitudes goes to my beloved father, Dr. Mohsen Mohebi, who thought me to be ambitious and has motivated me to pursue high-level educations. He admired me in successful moments of my life and though me not to give up in difficult situations.

I would like to thank my beloved mother, Ms. Zohreh Menbari for bestowing a calm and enthusiastic atmosphere for studying since I was a child. When I was at the last stage of my PhD, despite all the hurdles, she came to Canada and stayed with us for several months to provide a peaceful and gratified environment to finish my thesis.

My exceptional gratefulness goes to my beloved husband, Dr. Abbas Ahmadi, for his supportive and encouraging role in my life. Without his incentive and understanding it would have been impossible for me to finish this work. He is not only my best friend but a knowledgeable person whom I shared and discussed my PhD research with.

I would also like to thank the donors of Mik Pintar scholarship in MRI application and Dr. T.E. Unny Memorial Award for providing encouraging funding supports during my PhD study.

My deep appreciation goes to all my friends and colleges who helped me during my PhD study, specifically my best friends, Dr. Zohreh Azimifar, Dr. Mahtab Kamali, Dr. Atefeh Mashatan, Zahra Mousavian, Nasim Paryab and Bita Roushanei,

for their emotional support and unconditional friendship. I would like to thank Ms. Vicky Lawrence for her kind supports and my college, Ms. Ying Liu.

At last but not the least, I have been very grateful to have an angle from the heaven in the last year of my PhD, my beloved daughter, Saba. She brought enjoyable moments and happiness to my life during the most important phase of my PhD journey.

Dedication

To my parents

To my husband and my daughter

Contents

List of Tables

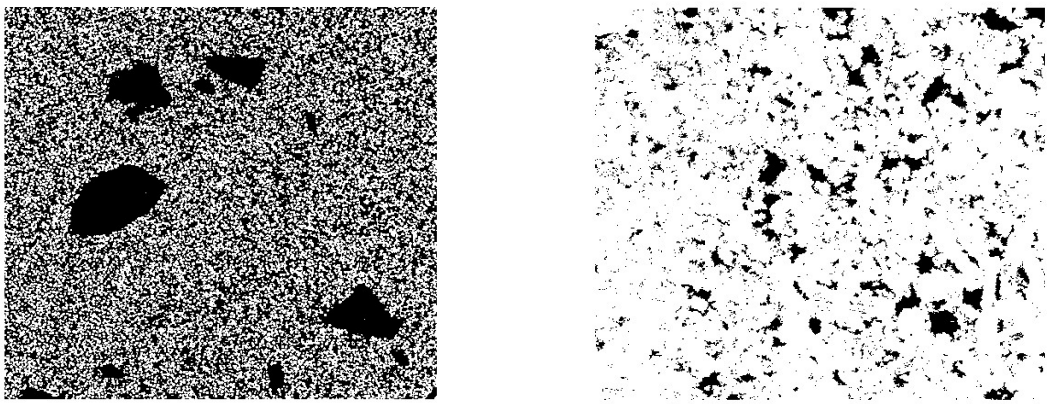
List of Figures

Chapter 1

Introduction

Scientific imaging plays a significant role in research, especially with the availability of sophisticated imaging tools, including magnetic resonance imaging (MRI), scanning electron microscopy, confocal microscopy, computer aided X-ray tomography, and ultrasound, to name only a few. Because of the significant research funding and public interest in medical imaging and remote sensing, these aspects of scientific imaging have seen considerable attention and success.

There is an enormous variety of imaging problems outside of medicine and remote sensing, where we would argue the current image processing research to be relatively rudimentary, and where substantial contributions remain to be made. One such area is that of porous media [1]. Porous media are the permeable materials such as cement, concrete, cartilage, bone, wood, and soil, with corresponding significance in the construction, medical, and environmental industries. Porous media possess complex, random-like structures which are difficult to describe quantitatively. However, they are not totally random, and there is a mixture of organization and randomness that can be described using probability theory, statistical models and random fields theory [1, 2, 3]. In analyzing the fluid transport, electromagnetic and mechanical properties of porous media, different realizations of porous media should be available. The 2D high resolution images of porous media (with resolution around $1.8\mu m/pixel$), shown in Fig. 1.1, obtained by microscopic imaging, provide valuable information on the large and small-scale structures of pores and solids. However, obtaining 2D images is invasive, requiring a series of processes such as cutting, polishing and exposure to air, which can affect the real properties of the porous materials. On the other hand, 3D samples of porous media can be obtained using MRI and X-ray computed tomography [4, 5]. However, the 3D samples are very low in resolution (in some cases around 100 $\mu m/pixel$) in comparison



Sintered Glass Spheres Carbonate Rock

Figure 1.1: Examples of 2D porous media images

with with the 2D samples, such that only some of the pore/solid structures can be resolved. In order to study different properties of porous materials such as fluid transport, permeability and conductivity, the real measured 2D or 3D samples are not sufficient and artificial samples are required to be generated in a porous media *reconstruction* task [3].

1.1 Motivation

Porous media reconstruction is aimed at generating *typical* artificial samples obeying certain statistical features. Therefore, the reconstruction task is a sampling process by which multiple samples are generated from a probability model. The model is defined based on different statistical descriptors such as chord-length, two-point correlation and histogram models [3, 6], using high resolution training data. Depending on the absence or presence of measurements, the process is defined as prior or posterior sampling [7]. The overwhelming majority of porous media simulation methods in the literature perform prior sampling, in which the reconstruction is solely based on the learned prior model [3, 6, 8, 9, 10, 11].

However, the growing availability and use of tomographic and MRI measurements means that in many cases it is desirable to study a particular physical sample, in which measurements are also available, and the reconstruction should reflect the measured sample. Thus, a *posterior sampling* approach [12] aimed at *fusing* the

information represented by the measurements and prior model, is required.

Fusing images at different resolutions and modalities has captured the interest of many researchers specially in the area of medical imaging and remote sensing, using super-resolution, multi-resolution, and image registration methods [13, 14, 15, 16]. Considering the existing methods in the literature, fusing the images of porous media is still a challenging task, due to the following reasons:

- Reconstruction of porous media is a sampling task by which typical samples of porous media with stochastic variability are generated from a statistical prior model , while the data fusion methods in the image processing literature are mainly aimed for finding an estimate.
- Despite many super-resolution and image registration problems, the porous media images at different resolutions, are not derived from the same scene, rather they are obtained from the same material, and consequently the images only share the same statistical features.
- In porous media reconstruction we do not have the luxury of having multiple low-resolution measurements from the same scene, as the case in some super-resolution problems, rather a few shots of each type of measurements are available.

In addition to the above reasons, most of the literature on multi-resolution image fusion, mainly based on wavelets, assumes continuous-state representation of the images – which is not the case with for porous media images. Therefore, a novel image fusion approach is required in order to generate artificial samples of porous media images.

1.2 Problem Summary

In order to reconstruct artificial samples of porous media a statistical data fusion approach is proposed in this thesis. The existing literature on reconstruction of porous media is mainly based on prior sampling in which artificial samples are generated from a statistical prior model and the low resolution measurements are not considered in the reconstruction process explicitly [3, 6, 17, 18]. The proposed statistical fusion approach in this thesis addresses fusing the information obtained from from low resolution measurements with the prior model, based on a Bayesian framework. The prior model is learned from the 2D high resolution training data

and the measurements are fused with the prior model using a measurement model. The artificial samples are then generated using a posterior sampling approach, in order to infer details at a higher resolution. Based on the proposed framework for data fusion, three main problems will be addressed in the thesis:

- **Two-scale porous media reconstruction**

  Low resolution measurements of porous media fail to resolve the pore/solid structures at multiple scales, rather the measurements are only able to resolve the large-scale structures while the fine/small-scale structures are unresolved. However, the fine/small-scale structures *are* constrained by the measurements, therefore there should be some possibility of inferring high resolution details from the pattern of low-resolution measurements. Using both low resolution measurements and a prior model learned from high resolution microscopic samples, a posterior sampling approach is proposed in Chapter 4 to reconstruct two-scale porous media images.

- **Annealing with constraints**

  Reconstruction of porous media has been studied broadly by sampling from the prior model using simulated annealing. Since porous media images possess complex structures at multiple scales, a single prior model can not reflect the variability of the structures at different scales. Using an additional information in the model as a constraint, a constrained sampling approach for porous media reconstruction is proposed in Chapter 5. The proposed approach is different from current porous media reconstruction methods in which the probability distribution is maximized by simulated annealing. Rather than maximizing the probability distribution, we sample from the constrained probability distribution using simulated annealing and Gibbs sampler.

- **Hierarchical posterior modeling**

  The reconstruction task becomes intractable when the size of the samples increases, since it is based on simulated annealing which is a slow convergence algorithm. Hierarchical sampling approaches for binary image synthesis have been applied to tackle this problem, in the case of sampling from the prior model. However, in the posterior sampling case, relating the measurement with an unknown at the coarse scale is still a challenging task. We propose a measurement model at each scale in Chapter 6, which defines how the measurement constrains the unknown at a given coarse scale.

Each of the mentioned problems are discussed separately with the proposed solutions in the thesis along with experimental results and evaluations.

1.3 Thesis Layout

The thesis consists of seven chapters including the introduction. Chapter 2 contains a brief overview of the background materials studied to pursue this research. In Chapter 3, the main framework proposed for image fusion is introduced. Chapter 4 contains the contribution on two-scale porous media reconstruction using different types of low resolution measurements. In Chapter 5, the constrained sampling approach using simulated annealing is proposed. The proposed approach is able to generate samples with more variable structures at different scales, as compared to the unconstrained sampling approach. In Chapter 6, a new hierarchical measurement model is proposed for the proposed hierarchical posterior sampling framework. Finally, in Chapter 7 concluding remarks and possible future research directions are provided.

Chapter 2

Background

This chapter contains a brief overview and description on some of the background materials used to pursue research for the thesis. The chapter starts with describing the forward and inverse problem and continues with a brief description of porous media. Then the image models used in this research are introduced and discussed followed by an overview of Monte Carlo Markov Chain sampling methods and simulated annealing. Finally, different hierarchical approaches used in image modeling and estimation are described.

2.1 Forward and Inverse Problems

In most physical systems two main sets of data are involved: a set of variables Z, which describe the state of the system and a set of data M, known as observations, indicating feasible measurements on that system [19].

Normally, Z is an ideal, complete representation of the system which refers to original data, whereas M is limited to the physics of the measuring device, so it is incomplete and irregular and does not provide detailed information on the system.

The task of inferring or computing measurements from the original data is known as *forward problem*. In this problem the goal is usually to compute M from Z according to some physical model f:

$$M = f(Z) \tag{2.1}$$

Forward problems are, *by definition*, categorized as easy problems, since the process of inferring M from Z can be fully characterized by a given physical model f, that is also referred to as *forward model*.

However in many cases, the original data are not known, and they should be determined given a set of measured values. In this case we are coping with another kind of problem called *inverse problem*. An inverse problem is to infer the original data from one or multiple observations of the data. In other words, solving an inverse problem involves finding the inverse of the forward model f in Eq. (2.1). Therefore, if the forward model can be represented mathematically, and is invertible, possibly we can find the solution for the inverse problem as follows:

$$Z = f^{-1}(M) \tag{2.2}$$

However, it is usually hardly possible to determine Z from f^{-1}, because [20, 21]:

- In many problems, f^{-1} may not even exist, for example when the measurements are noisy, we can not define an explicit and exact relationship between Z and M. In this case we say that the *existence* is failed.
- Measurements are usually incomplete and they are not enough to reconstruct a complete original set of data. In many cases, different sets of original data lead to the same set of measurement data, such that a unique Z can not be inferred from M. In other words the *uniqueness* is failed.
- In finding Z from f^{-1} we may cope with an unstable situation in which a small change in the measurement may lead to a huge change in Z. In this case we say that *continuity* is failed.

According to Hadamard definition [22], a problem is well-posed if the existence, uniqueness and continuity conditions hold, otherwise it is ill-posed. An ill-posed problem turns to a well-posed problem by adding additional information using regularization constraints or prior knowledge [21, 22].

2.1.1 Bayesian and non-Bayesian estimation

Many problems in image processing are ill-posed, such as image reconstruction, image de-noising and image enhancement. However, an ill-posed problem can become well-posed with regularization methods by adding constraints or a priori knowledge. For instance, in the case that the existence fails, the inverse problem can be modified to finding the closest Z, instead of finding the exact data from which the measurement M is derived, i.e.

$$\hat{Z} = arg\min_{Z} \|f(Z) - M\| \tag{2.3}$$

Here, we are looking for an *estimate* $\hat{Z}$ of Z satisfying the minimization constraint. Therefore, by regularizing the problem we are solving an estimation problem in which an approximation of the original data is derived. Usually the estimated data is consistent with the measurement (to some extend) and satisfy the constraint. Depending on the nature of the problem and/or the degree to which we believe in the constraint or the priori knowledge, an inverse problem can be regularized by a deterministic constraint or by a stochastic prior model.

In the deterministic (non-Bayesian) approaches, Z is just an unknown that we need to estimate. In such cases, estimation approaches such as the usual least squares formulation [21], maximum likelihood (ML) estimation or smoothing constraints can be applied. For instance, the ML estimator is obtained by maximizing the likelihood function

$$\hat{Z} = arg\max_{Z} p(M|Z) \tag{2.4}$$

where $p(M|Z)$ is also known as the measurement model.

In the Bayesian approach [23], we consider that the original data Z are a set of random variables obeying some statistics known as prior model. In this case we believe that the stochastic behavior of Z is an inherent part of the problem. Then Bayesian estimation approaches, such as Maximum A Posteriori (MAP) or Bayesian least square estimation [21], are usually applied to solve the inverse problem. For instance, the MAP estimator which is a Bayesian version of the ML estimator, is obtained by maximizing the posterior probability distribution,

$$\hat{Z} = arg\max_{Z}\ p(Z|M) = arg\max_{Z} p(M|Z)p(Z) \tag{2.5}$$

where $p(Z)$ is the prior model describing the statistical behavior of Z and $p(M|Z)$ is the likelihood function. The above equality is obtained via Bayes' rule [23].

2.1.2 Sampling

Estimation methods usually find *optimum* solutions which maximize/minimize a pre-defined set of criteria. However in some applications, such as texture analysis and porous media reconstruction, we are not interested in finding the best/optimum solution, rather we seek a random realization of the original data meeting a given set of criteria (prior model) to some extent while being consistent with the measurements [21]. The problem of generating such realizations or *samples* is referred to as sampling. Depending on the absence/presence of measurements in a problem, we have prior/posterior sampling, described as follows:

- *Prior Sampling*: given a prior probability model $p(\cdot)$ of the original data, prior sampling means generating random samples from the prior model. Prior sampling is interesting in
  - Studying prior models: since a prior sample depends only on the prior model, it can reflect insights regarding model strengths and weaknesses.
  - Image analysis and reconstructions: prior samples are applicable in texture generation and porous media reconstruction, to synthesize and study different aspects of the data [6].
- *Posterior Sampling*: given a prior probability model $p(\cdot)$ for the original data and a set of measurements M, posterior sampling is generating samples from the posterior distribution $p(\cdot|M)$. If $\hat{Z}$ denotes the solution of the estimation problem, then based on Eq. (2.3) the posterior sample $Z_{|M}$ is

$$Z_{|M} = \hat{Z} + \tilde{Z}_{|M} \tag{2.6}$$

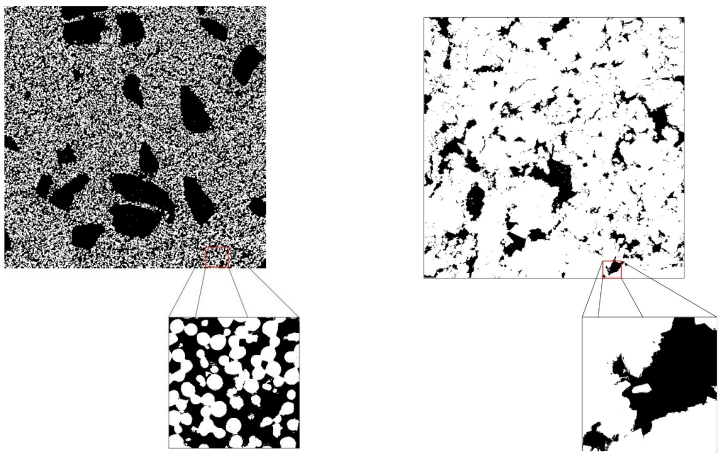
  where $\tilde{Z}_{|M}$ is the estimation error. In fact, a typical posterior sample has aspects of measurements, inferred by $\hat{Z}$, and also the prior model, inferred by $\tilde{Z}_{|M}$. [7].

  As in estimation problems, two types of information are used in posterior sampling, i.e. measurement and prior model. Therefore, there is a possibility to cover the gap due to any incompleteness or inconsistency in one with another. In estimation, we seek the optimum solution and it does not represent a typical or representative sample of the system being studied. However, a posterior sample is a *typical* random realization which is consistent with the model and the measurements.

In this research a posterior sampling approach is considered in order to generate artificial samples of scientific images such as porous media images.

2.2 Porous Media

A practical and important class of scientific images are the 2D/3D images obtained from porous materials such as concretes, soils, sandstone, active carbon and glass. Fig. 2.1 shows two examples of 2D images of porous media. Most porous media possess a chaotic structure that is difficult to describe qualitatively [1, 3], but they are not totally random: there is a mixture of organization and randomness that



(a) Sintered Glass Spheres (b) Carbonate Rock

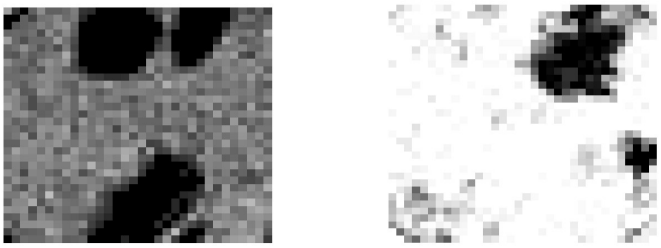
Figure 2.1: High resolution samples of porous media. The samples have different structures at different scales.

makes them difficult to characterize and study. A porous material contains multi-phases, mainly limited to two: void and solid. The pore or void phase shown as the black areas in Fig. 2.1, can transport fluids/ gas and the rest shown as white in Fig. 2.1 are solid.

2.2.1 Reconstruction of porous media

In order to study the permeability, conductivity, and transport properties of porous materials, 2D/3D high resolution samples are required. But obtaining high resolution data from a physical material usually requires cutting, polishing and exposure to air, all of which affect the properties of the sample. On the other hand, 2D/3D samples from Magnetic Resonance Imaging (MRI) can be obtained in-situ, with no modification of the sample, however the measurements are noisy and taken at a very low resolution, such that only comparatively large pores are resolved (as shown in Fig. 2.2). Therefore, high resolution samples of porous media need to be generated through a *reconstruction* process [3].

Porous media 2D/3D reconstruction involves combining certain features of a set of real measured media and generating artificial media with the same feature, usually using numerical methods. The features can be represented by a statistical model, specific structure, and/or even the real measured data.



(a) Sintered Glass Spheres (b) Carbonate Rock

Figure 2.2: 2D slices of MRI measurements [4] obtained from the same samples shown in Fig. 2.1. The measurements in (a) and (b) and their corresponding images in Fig. 2.1 are not from the same scene, rather they are obtained from the same material. The measurements only resolve the large scale structures.

One of the approaches used for porous media reconstruction is to use Gaussian linear and non-linear filters [1, 24]. The Gaussian filtering method is limited to isotropic media and is formulated based on porosity and two-point correlation function information. In addition, this approach is solely based on Gaussian random field which is not able to model non-stationarity and the multi-scale structures presented in porous materials [1, 9]. Therefore, the Gaussian filtering method is of limited use in porous media reconstruction.

Another approach, which is the main focus of this research, is based on defining statistical models describing the characteristics and different structures of pore and solid [3] and using computational methods to assert the same statistical features represented by the prior model to generate an artificial sample. The parameters of the model are usually learned from high resolution training samples. The statistical models can be considered as a prior probability distribution describing the probability of having a certain structure or feature in a porous medium. The artificial samples are then generated by sampling from the prior model [9, 6, 3]. Most of these methods are based on creating a single, fixed, target distribution (e.g. correlation function or chord-length distribution). However, realistic samples of porous media have complicated structures, so there is a need to have an average model containing different aspects of each single distribution. Although some research has been done recently on incorporating the measurements information such as the overall porosity and the information on the large-scale structures in the reconstruction task [5, 25], they do not couple the measurements with the prior model explicitly.

In the next section a brief overview of some of the models used to describe the statistical behavior of porous media is provided.

2.3 Image Modeling

Bayesian image analysis involves representing image attributes as a prior model. The prior model can be defined based on the local or non-local interactions between different pixels in an image. In this section we review three local models including Ising, Potts and histogram, and two non-local models – two-point correlation and chord-length.

While simple image models can be obtained from image statistics such as mean, variance, histogram and correlation [26, 27] a more general approach is to use random fields, probably the most common class of image models. A random field is a two-dimensional extension of a random process, and can be defined as follows:

Definition (random field): Let Z be a finite set of sites or locations $Z = \{z_1, z_2, ..., z_N\}$ in a lattice, where each $z_i \in Z$ can have finite possible states from Λ. Then Z is a configuration such that

$$Z = \{ z_i \mid 1 \leq i \leq N\} \tag{2.7}$$

We can consider a probability measure or distribution p on the set of all configurations, Ω, such that

$$p(Z) \geq 0 \; and \; \sum_{Z \in \Omega} p(Z) = 1. \tag{2.8}$$

In this case Z is a random field. An $(n \times m)$ image can also be represented as a random field as shown in Eq. (2.7). Each site in image Z can be represented as $z_{i,j}$, where $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

2.3.1 Markov random fields

There are variety of models proposed for random fields analysis. Markov random field (MRF) models [28] consider conditional decoupling of a set of sites from the rest of the field. In image processing and modeling problems the focus is more on statistical modeling of images with specific emphasis on MRF, [26, 28]. Some of the reasons for using MRF are [29] as follows:

- Applying spatial context in image classification deals with assigning a symbolic label to a group of pixels. In these problems, the contextual information is taken locally, an MRF can provide a flexible mechanism to model spatial connection and correlation between the local information.

- In a random field, if we want to consider the probability of a site taking a specific value, we need to study and consider the rest of the field which is computationally very expensive. MRF models are based on considering a local neighborhood rather than the whole field to compute the aforementioned probability.

MRFs are characterized by conditional dependency between a set of pixels with its neighbors. The neighborhood structure is defined as follows.:

Definition (neighborhood structure): A set $\mathcal{N}(i,j)$ is said to be neighbors of a site or a pixel $z_{i,j}$ in an $n \times m$ lattice if it contains all the neighbors of that site based on a neighborhood relationship. The neighboring relationship has the following properties:

- a site is not neighboring to itself $z_{i,j} \notin \mathcal{N}(i,j)$
- the neighboring relationship is mutual $z_{i,j} \in \mathcal{N}(i',j') \iff z_{i',j'} \in \mathcal{N}(i,j)$

For a regular lattice , $\mathcal{N}(i,j)$ is defined as the set of nearby sites within a radius r, based on the Euclidean distance [30]

$$\mathcal{N}(i,j) = \{z_{i',j'} | (i-i')^2 + (j-j')^2 \leq r^2,\ i' \neq i,\ j' \neq j\} \tag{2.9}$$

where r takes an integer value. Note that sites at or near the boundaries have fewer neighbors. Fig. 2.3 (a) and (b) shows the first and second order neighborhood structures, respectively and in (c) higher order neighborhood structures are illustrated as well. The kth order neighborhood structure encompasses all the elements in of the lower order neighborhood structures.

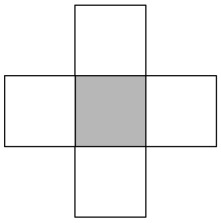
By changing the lattice of sites in a lexicographic order the columns of the lattice are stacks from left to right in a single array. Therefore, we can re-write image Z as Eq. (2.7) and we can represent the location of a site in the image with a single index i. For the sake of simplicity, we consider the lexicographic order and consequently call the neighborhood set of site z_i as $\mathcal{N}(i)$.

Definition (Markov random field): A random field Z with probability distribution p is Markov if :

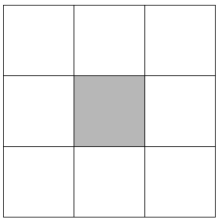
$$p(\ z_i = z^*\ |\ Z_{/\{i\}})\ =\ p(\ z_i = z^*\ |\ \mathcal{N}(i)). \tag{2.10}$$

and consequently

$$p(Z = Z^*)\ =\ \prod_{i=1}^{n} p(z_i = z_i^* | \mathcal{N}(i)) \tag{2.11}$$



(a) First-order neighborhood structure



(b) Second-order neighborhood structure

5	4	3	4	5
4	2	1	2	4
3	1		1	3
4	2	1	2	4
5	4	3	4	5

(c) Neighborhood structures at different orders

Figure 2.3: Two examples of neighborhood structure, in (a) and (b). The numbers in (c) denotes the order of the neighborhood structures, and the location of a certain number in (c) shows the extent of the neighborhood structure.

where $Z_{/\{i\}}$ excludes the site i form the whole random field Z. MRFs can be used for modeling real microstructural/heterogeneous media whose properties vary randomly and smoothly in space [2].

2.3.2 Gibbs random field

Gibbs Random Fields (GRFs) were originally used in statistical physics to study the thermodynamic characteristics of interacting neighboring particles in a system [31]. They are used as models for equilibrium states of large physical system. For these fields a probability distribution function called Gibbs distribution is defined as

$$p(Z) = \frac{e^{-\frac{H(Z)}{T}}}{\mathcal{Z}} \tag{2.12}$$

where $H(Z)$ is an *energy function* capturing the interaction between neighboring particles, T is a parameter usually called the temperature, and $\mathcal{Z}$ is a normalization factor, such that

$$\mathcal{Z} = \sum_{Z \in \Omega} H(Z). \tag{2.13}$$

The most important part of the Gibbs distribution is the energy function $H(Z)$. In fact, this function reflects the whole characteristics of the model. In general the energy function $H(Z)$ can be written as

$$H(Z) = \sum_{c \in C} V_c(Z) \tag{2.14}$$

where c is a clique—a single site or a set of sites that interact with each other— and C denotes the set of all cliques. The potential function V_c represents the interaction between sites in clique c [31].

According to the Hammesley-Clifferd [29] theorem, GRFs and MRFs are equivalent. Both MRFs and GRFs are used to quantify the spatial interactions of observed values at the sites of a field and give a probability for any configuration of that field. The algorithms that exist for analyzing MRFs usually assume that the field is Gaussian, where GRFs' algorithms can also handle non-Gaussian, non-linear problems [13]. Moreover, Gibbs probability distribution is used for the discrete-state random fields [32, 31]. In this research we are coping with the reconstruction of discrete-state random fields (porous media images are binary), therefore, we consider Gibbs probability distribution to model the characteristics of porous media images.

2.3.3 Classic models

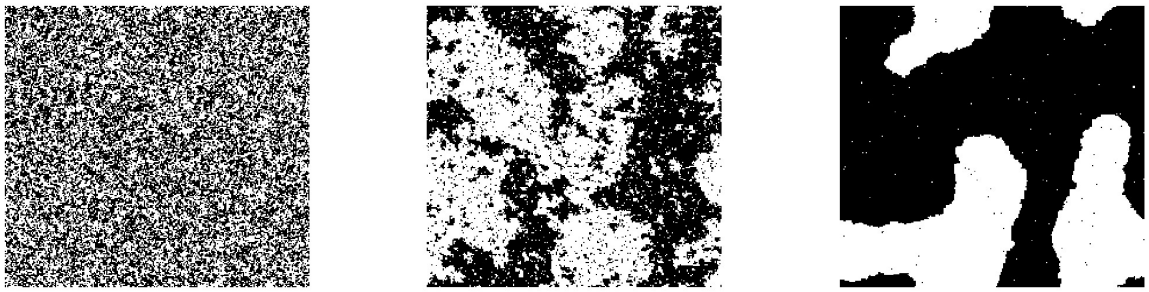
In this section three image prior models are discussed: Ising, Potts and histogram. The models are pixel-based and the interaction between pixels are local. The binary Ising model [31], based on defining a first-order neighborhood for a pixel, is considered due to its simplicity and for comparison purposes. The Potts model [31] as a generalization of the binary Ising model is also discussed for the case that the image is not binary. The histogram model [6], a pixel-based and local model, is a non-parametric model in which the neighborhood size is flexible.

Ising model

The Ising model was first introduced in quantum physics where each field element reflects the discrete quantum state of an atom [21]. Each site can take two possible values, 0, and 1, and a first-order neighborhood structure, as in Fig. 2.3 (a), is considered for each site, i.e. for every site $z_{i,j}$ in the field, sites $z_{i,j+1},\ z_{i,j-1},\ z_{i+1,j},\ z_{i-1,j}$ are defined as its neighbors. The general form of the energy function for an Ising model is

$$H(Z) = a \sum_{i,j} z_{i,j}\ (z_{i,j+1} + z_{i,j-1} + z_{i+1,j} + z_{i-1,j}) + b \sum_{i,j} z_{i,j}\ h_{i,j} \qquad (2.15)$$

where a and b are model parameters and $h_{i,j}$s are constants related to an external field constraint. Here a clique is either a site $(z_{i,j})$, or two neighboring sites $(\{z_{i,j}, z_{i,j+1}\},\ \{z_{i,j}, z_{i,j-1}\},\ \{z_{i,j}, z_{i+1,j}\},\ \{z_{i,j}, z_{i-1,j}\})$. The random field shown in Fig. 2.4 is generated using Ising model at different values of temperature T.



(a) $T = 5$, "hot" (b) $T = 2.27$, "warm", (c) $T = 1.25$, "cold"

Figure 2.4: Samples from Ising model at different temperatures

Potts model

The Potts model [31] originally used in statistical physics, is the generalization of the Ising model in which each site can take q possible values. It is sometimes referred to as q-Potts model.

This model identifies boundaries in the random field using the *Kronecker* function $\delta(a, b)$ defined as

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y. \end{cases}$$

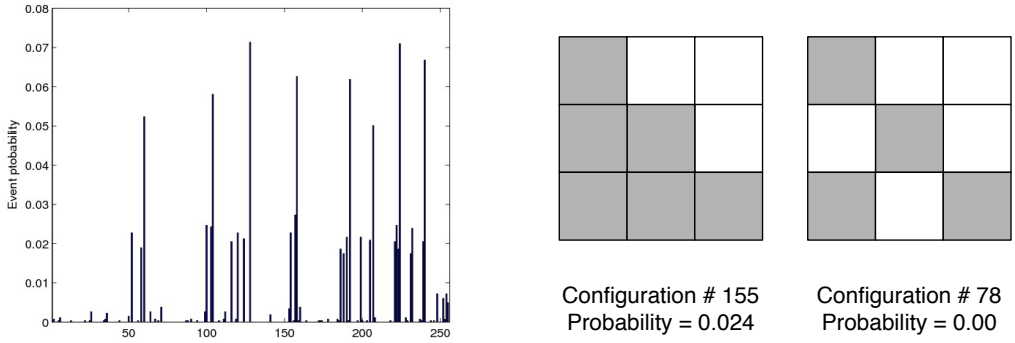
For a binary image, the Kronecker function measures the length of the boundary in the whole image. The energy function for the standard Potts model is defined as

$$H(Z) = a \sum_{i,j} [1 - \delta(z_{i,j}, z_{i,j-1})] + [1 - \delta(z_{i,j}, z_{i-1,j})] \sum_{i,j} z_{i,j}\ h_{i,j} \tag{2.16}$$

where a and b are model parameter and $h_{i,j}$s are constants related to external field interactions.

Histogram model

The histogram model [6] is non-parametric, keeping the entire joint probability distribution of a local set of pixels within a neighborhood. Choosing eight adjacent pixels as the neighborhood structure (as Fig. 2.3 (b)) leads to a non-parametric model containing a histogram of $2^9 = 512$ probabilities (for binary random field, each site can take two possible values). Fig. 2.5 shows such a normalized histogram, learned from the image in Fig. 2.1(a).



(a) Probability of 256 possible configurations

(b) Examples of two configurations and their probability

Figure 2.5: An example of probability distribution of $2^9 = 512$ configurations with black pixels at the center (a), with examples of two configurations and their probabilities (b). The random field is modelled non-parametrically via the probability of every configuration of the eight binary pixels surrounding a central pixel.

2.3.4 Porous media models

The reconstruction of porous media, which is the core objective of this research needs defining models describing their statistical properties and microstructures. The porous media models which are microstructural descriptors are defined based on structure/property relations [3]. Here, we introduce two types of porous media models: two-point correlation and chord-length model.

Two-point correlation model

The random space of a binary porous medium can be partitioned into two phases made up of pore and solid phases. Having two phases, the two-points correlation model [9, 3] considers the probability of finding two vectors at two different specific positions $\vec{r}_1$ and $\vec{r}_2$, in the same phase.

The model defines an autocorrelation function

$$S_2^{(j)}(\vec{r}_1, \vec{r}_2) = E[\ I^{(j)}(\vec{r}_1)\ I^{(j)}(\vec{r}_2)\ ] \tag{2.17}$$

where $E[\cdot]$ stands for statistical expectation and $I^{(j)}(r)$ is the characteristic function

$$I^{(j)}(r) = \begin{cases} 1 & \text{when } r \text{ is in phase } j \\ 0 & \text{otherwise.} \end{cases}$$

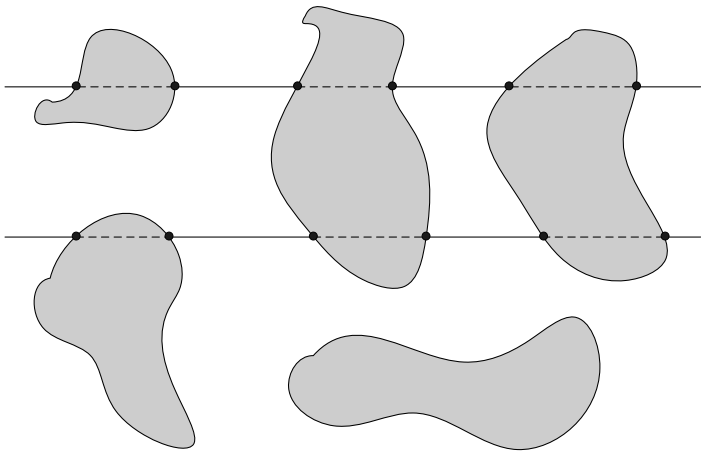


Figure 2.6: Chords are line segments between the intersection of two phases with an infinitely long line

For statistically isotropic media, $S_2^{(j)}(\vec{r}_1, \vec{r}_2)$ depends only on the distance $r = |\vec{r}_1 - \vec{r}_2|$, between two points, and the autocorrelation function can be expressed as $S_2^{(j)}(r)$. If ϕ_j is the volume fraction of phase j, then [3]

$$S_2^{(j)}(0) = \phi_j \quad and \quad \lim_{r\to\infty} S_2^{(j)}(r) = \phi_j^2. \tag{2.18}$$

The two-point correlation function can be extended to n-point correlation function by considering the probability that n points at positions $r_1, r_2, \ldots, r_n$ are found in the same phase. Although extending to an n-point correlation function can give a better understanding of the porous media, the storage of the information and also the large computational time limit the use of high order ($n > 2$) correlation functions in practice.

Chord-length model

In this model a *chord-length density function* $p^{(i)}(c)$ is defined for phase i of the random field [3]. Chords are all of the line segments between intersections of an infinitely long line thrown in a two-phase random field.

The chord-length distribution function can be defined in terms of *Lineal-path function* $L^{(i)}(c)$. For statistical isotropic media, $L^{(i)}(c)$ is the probability that a

line segment of length c lies wholly in phase i when randomly thrown into the field. $L^{(i)}(c)$ can be obtained by counting the relative number of times that a line segment of length c is wholly in phase i.

The first derivative of the lineal-path function is related to the *cumulative distribution function* $F(c)$ associated with $p^i(c)$ and the second derivative is the chord-length distribution function. Therefore [3]

$$p^{(i)}(c) = \frac{\ell}{\phi_i} \frac{\sigma^2 L^{(i)}(c)}{\sigma c^2}. \tag{2.19}$$

where $\ell = \int_0^\infty y\ p^{(i)}(y)\ dy$.

The Chord-length distribution function can be defined for either phase in an image for lines at different orientations. It is common to limit the orientation to the horizontal and vertical directions in 2D. Defining this function only for one phase gives limited information about the other. Therefore, having this function for both phases (pore and solid) leads to better understanding of the image. For the two phase porous media, if Chord-length distribution function is defined for both phases, then we have a *dual Chord-length* model.

2.4 Monte Carlo Markov Chain Sampling Methods

Given a random field model defined for an image, the analyzing task such as sampling, restoration and segmentation is usually intractable analytically since the discrete configuration space on which the random field exists is very large; even for a binary $n \times n$ image the space has 2^{n^2} elements. More specifically, computing the normalization factor in the Gibbs probability distribution defined in Eq. (2.13), which should encompass all the 2^{n^2} configurations, is not possible, and consequently the probability distribution of the image can not be computed directly. Monte Carlo methods are useful in image analysis problems since they can simulate the random field regardless of the size of the image and optimize energy functions without computing the normalization factor for the whole configuration space.

Computing the expected value of a random field based on a probability distribution yields the most probable realization. Suppose that we want to compute the expected value of $\phi(Z)$ for a given function ϕ and a random field Z. Given the

probability distribution of Z as $p(\cdot)$, the formula

$$E[\phi(Z)] = \int \phi(Z)\ p(Z)\ dZ \tag{2.20}$$

can be used. However, computing this integral is hardly possible. Using numerical integration and the strong law of large numbers, the expectation can be estimated by generating n samples $(Z_1, \ldots, Z_n)$ from the probability distribution p and computing the ensembles, i.e.

$$E[\phi(Z)] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \phi(Z_i). \tag{2.21}$$

This method, known as Monte Carlo, is applicable when [33]

1. the probability distribution p can not be computed for a given Z, e.g. when p is Gibbs distribution and the normalization factor $\mathcal{Z}$ defined in Eq. (2.13) can not be computed,
2. the configuration space of Z is very large and sampling from this space is not possible.

In sampling problems, where large random fields are involved, generating samples from the probability distribution is still a problem. Monte Carlo Markov Chain (MCMC) methods based on constructing an ergodic Markov chain generate samples from the probability distribution. Running the Markov chain for a long time, MCMC methods provide samples which are statistically consistent with the original random field [31, 33].

MCMC have three main properties [31]:

1. A given configuration is updated in sequential steps,
2. Each step is governed by some probabilistic rule
3. The rule depends only on the number of the step and on the present configuration.

Gibbs and Metropolis-Hastings samplers are two main MCMC methods which are used in different image analysis applications. In the following sections, these methods are discussed briefly.

2.4.1 Gibbs sampler

The Gibbs sampler [34] algorithm is a technique for generating random variables from a distribution *indirectly*, without having to calculate the density [33]. In particular, it can generate samples from the Gibbs probability distribution defined in Eq. (2.12). The algorithm generates a Markov chain whose elements are sequential random fields, $Z^{(0)}$, $Z^{(1)}$, ,..., such that $Z^{(k-1)}$ and $Z^{(k)}$ can differ in at most one pixel. For the configuration space Ω, this iterative algorithm starts with an arbitrary $Z^{(0)} \in \Omega$. If the random field is written in lexicographic form, then given a site visiting schedule at every iteration, for the jth site, the chain proceeds from $Z^{(k-1)}$ to $Z^{(k)}$ as follows [29]:

$$Z_i^{(k)} = \begin{cases} Z_i^{(k-1)} & \text{if } i \neq j \\ \xi & \text{if } i = j \end{cases} \qquad for\ all\ i.$$

where ξ is a random variable with probability distribution

$$P(\xi = \lambda) = p(Z_i = \lambda \mid \mathcal{N}(i)). \tag{2.22}$$

This conditional probability distribution determines the likelihood of random variable Z_i to be equal to λ, given its neighbors. Using Gibbs probability distribution, we have

$$p(Z(i) = \lambda \mid \mathcal{N}(i)) = \frac{1}{\mathcal{Z}_i}\ e^{-\frac{1}{T}\ H(Z(i)=\lambda|\mathcal{N}(i))} \tag{2.23}$$

and

$$\mathcal{Z}_i = \sum_{\gamma \in \Lambda}\ e^{-\frac{1}{T}\ H(Z_i=\gamma|\mathcal{N}(i))} \tag{2.24}$$

The normalization factor, $\mathcal{Z}_i$ sums over all possible values from the state space Ω that Z_i can take. In particular, when the field is binary, ($\Lambda = \{0, 1\}$), Eq. (2.24) is simplified to

$$\mathcal{Z}_i = e^{-\frac{1}{T}\ H(Z_i=0|\mathcal{N}(i))} + e^{-\frac{1}{T}\ H(Z_i=1|\mathcal{N}(i))} \tag{2.25}$$

and consequently for $\lambda = 0$, Eq. (2.23) is simplified to

$$p(Z_i = 0 \mid \mathcal{N}(i)) = \frac{1}{1 + e^{-\frac{1}{T}\ (\ H(Z_i=1|\mathcal{N}(i)) - H(Z_i=0|\mathcal{N}(i))\ )}} \tag{2.26}$$

The updating process from $Z^{(k-1)}$ to $Z^{(k)}$ is repeated until we get independent samples [33] to compute the ensemble in Eq. (2.21). Algorithm (1) shows a summary of steps taken by Gibbs sampler to generate a meaningful sample from the Gibbs

probability distribution for a binary image. The Gibbs sampler and other types of MCMC methods are converged when a stationary chain of processes are generated. Several convergence diagnosis techniques are available usually based on studying the covariance of the samples to see how much they are independent from each other and determine whether the resulting chain is stationary or not [33, 35, 36].

Algorithm 1 Gibbs Sampler for a binary random field

1: Start with an initial field $Z^{(0)}$ by assigning 0 or 1 to each site randomly.
2: $k \leftarrow 0$
3: Re-order elements of $Z^{(k)}$ according to a site-visiting schedule
4: **repeat**
5: **for** $i = 1$ to N **do**
6: $Z_i^{(k)} \leftarrow 1 - Z_i^{(k)}$ {change the value of pixel i}
7: **if** $Z^{(k)}(i) = 0$ **then**
8: $p \leftarrow \frac{1}{1+exp(-\frac{1}{T}\ (\ H(Z_{(k)}(i)=1|\mathcal{N}_i)-H(Z_{(k)}(i)=0|\mathcal{N}_i)))}$
9: **else**
10: $p \leftarrow \frac{1}{1+exp(-\frac{1}{T}\ (\ -H(Z_{(k)}(i)=1|\mathcal{N}_i)+H(Z_{(k)}(i)=0|\mathcal{N}_i)))}$
11: **end if**
12: Pick a point $r \sim U(0,1)$
13: **if** $p < r$ **then**
14: $Z_i^{(k)} \leftarrow 1 - Z_i^{(k)}$ {undo the change}
15: **end if**
16: **end for** {$Z^{(k+1)}$ is generated}
17: $Z^{(k+1)} \leftarrow Z^{(k)}$
18: $k \leftarrow k + 1$
19: **until** Convergence condition is met

The Gibbs sampler is not practical when the number of possible states, $|\Lambda|$, is large, since for each site Z_i, $e^{-\frac{1}{T}\ H(Z_i=\lambda|\mathcal{N}(i))}$ needs to be computed for all $\lambda \in \Lambda$.

2.4.2 Metropolis-Hastings sampler

The Metropolis-Hasting sampler [37] is a mechanism for generating a Markov chain $Z^{(1)}$, $Z^{(2)}$, ... given an arbitrary starting point $Z_{(0)}$. The chain proceeds from $Z^{(k-1)} = x$ to $Z^{(k)} = w$ by suggesting a change according to a *proposal distribution* $q(\cdot|Z_{(k-1)} = x)$. The proposal distribution may be chosen to be random, although depending on the nature of the problem, there may be efficiency advantages in one

form over another [33, 38]. The candidate point w is accepted with the following probability

$$p(Z^{(k)} = w|Z^{(k-1)} = x) = min\left\{\frac{p(Z^{(k)} = w)}{p(Z^{(k-1)} = x)}\ \frac{q(Z^{(k-1)} = x|Z^{(k)} = w)}{q(Z^{(k)} = w|Z^{(k-1)} = x)}\ ,\ 1\right\}. \tag{2.27}$$

If the suggestion is accepted, then $Z^{(k)} = w$ otherwise $Z_{(k)} = x$. In other words, in the Metropolis-Hasting method, depending on x, a new element w is proposed by some procedure. If w is better than x in terms of energy function, the change is accepted, if not, then it still has a chance to be accepted. The key advantage of this method is that there is no need to compute the normalization factor in the probability distribution p since it will be cancelled out in Eq. (2.27), therefore in contrast with Gibbs sampler, this method is also applicable in the case that $|\Lambda|$ is large.

The same as Gibbs sampler, the process can be repeated till a relatively meaningful number of samples are generated [33]. Algorithm (2) shows a summary of steps taken by Metropolis-Hasting sampler to generate a meaningful sample from the Gibbs probability distribution for a binary image. The proposal distribution q is considered to be a uniformly random scheme.

2.4.3 Simulated annealing

Simulated annealing (SA) is an optimization method proposed by Kirkpatrick et al.(1983) [39]. It finds the global minimum/maximum of a function that may possess several local minima/maxima [40]. In image analysis problems, SA usually seeks the maximal modes of the probability distribution on image space. For a Gibbs random field, MCMC sampling methods provide a random sample of the Gibbs distribution. However, we can apply SA to find the *most probable* sample, by maximizing the Gibbs probability distribution which can be done by decreasing the temperature (T) during the sampling process. For that purpose, an initial configuration of the random field is picked randomly and one sweep is done with the Gibbs or Metropolis-Hasting sampler at a specific starting temperature T_0. Decreasing T_0 to T_1, the next sweep is done with T_1 and the process will continue while $T_k \to 0$. As T is decreased to zero, the Gibbs field converges to the uniform distribution on the space of maximal modes which correspond to the field maximizing the Gibbs probability distribution – minimizing the energy function.

Let $H_0(Z) = H(Z) - \min(H(Z))$, and for simplicity let $\min(H(Z)) = 0$, hence we have [31]

Algorithm 2 Metropolis-Hasting Sampler for a binary random field with uniform random proposal distribution

Start with an initial field $Z^{(0)}$ by assigning 0 or 1 to each site randomly.
$k \leftarrow 0$
Re-order elements of $Z^{(k)}$ according to a site-visiting schedule
repeat
    for $i = 1$ to N **do**
        Propose a value for $Z_i^{(k)}$ as α, randomly. $\{\alpha \in \{0, 1\}\}$
        $Pr \leftarrow \frac{p(Z_i^{(k)}=\alpha)}{p(Z_i^{(k)}=1-\alpha)}$
        if $Pr \geq 1$ **then**
            $Z_i^{(k)} \leftarrow \alpha$
        else
            Pick a point $r \sim U(0, 1)$
            if $Pr \geq r$ **then**
                $Z_i^{(k)} \leftarrow 1 - \alpha$
            end if
        end if
    end for {a sample random field $Z^{(k+1)}$ is generated}
    $Z^{(k+1)} \leftarrow Z^{(k)}$
    $k \leftarrow k + 1$
until Convergence condition is met

$$
\begin{aligned}
p(Z) &= \frac{\exp(-\frac{1}{T}\ H(Z))}{\sum_{Z\in\Omega} \exp(-\frac{1}{T}\ H(Z))} \\
&= \frac{\exp(-\frac{1}{T}\ H_0(Z))}{\sum_{H(Z)=0} \exp(-\frac{1}{T}\ H_0(Z)) + \sum_{H(Z)>0} \exp(-\frac{1}{T}\ H(Z))} \\
&= \frac{\exp(-\frac{1}{T}\ H_0(Z))}{|M| + \sum_{H(Z)>0} \exp(-\frac{1}{T}\ H(Z))} \xrightarrow{T\to 0} \begin{cases} \frac{1}{|M|} & \text{if } H(Z) \text{ is minimal} \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

Therefore, as T decreases the probability distribution tends to a uniform probability distribution on the space of all configurations that minimize H. Consequently, applying Gibbs or Metropolis-Hasting sampler on this space leads to a random field with maximum likelihood.

The parameter T is called the *temperature* [31]. In fact, the idea of SA comes from physics, where large physical systems tend to states of minimal energy, called ground states, if cooled down slowly enough. If T decreases rapidly, then the particles of the system do not have enough time to interact to achieve the equilibrium point of the system, and the process will stick into local minima or non-equilibrium point [34]. According to [34] the optimum annealing schedule for T is the logarithmic schedule

$$T_k \propto 1/log(1+k) \tag{2.28}$$

However, the logarithmic schedule is not really practical and there are other types of annealing schedules such as *exponential* schedule, which are more practical [6]

$$T_k = T_0 \; b^k \tag{2.29}$$

where a is constant and b is sufficiently close to 1 ($b = 1 - \epsilon$, for small $\epsilon > 0$) to insure slowly enough decreasing temperature. The exponential annealing schedule does not guarantee obtaining the global optimum at the end rather it provide a local optimum, very close to the global solution.

Algorithm (3) shows a summary of steps taken by SA for exponential annealing schedule to generate the most probable random field of the Gibbs probability distribution, for a binary image. The practical convergence condition for the algorithm is one of the following cases

- when T is smaller than a specific value
- when the energy function $H(Z)$ reaches a pre-defined small value,
- when the energy function $H(Z)$ settles down, i.e. it does not change for a long range of time.

Where in this thesis the last criteria has been considered.

Algorithm 3 Simulated Annealing with exponential annealing schedule

1: Choose suitable T_0 and b
2: $k \leftarrow 0$
3: **repeat**
4: Do the Gibbs/Metropolis-Hasting Algorithm for $Z^{(k)}$ and T_k
5: $k \leftarrow k + 1$
6: $T_k = T_0\ b^k$
7: $Z^{(k)} \leftarrow Z^{(k-1)}$
8: **until** convergence condition is satisfied

2.5 Hierarchical Modeling of Markov Random Fields

The energy-based models for image analysis and reconstruction which are described in Section 3.2 use the deterministic or stochastic iterative algorithms such as Gibbs sampler and Metropolis Hasting sampler along with simulated annealing. Such iterative methods are very slow in convergence and cause high computational time [41]. The hierarchical approaches such as multi-grid, multi-scale and multi-resolution algorithms [41, 42, 43], are able to decrease the computational time significantly while improving the quality of the results. An important class of hierarchical approaches are Multi-grid method. However, multi-grid method is suitable for continuous-state random fields and is based on incremental refinement of the result at each scale. Since porous media images are binary, we review different types of energy-based hierarchical approaches which can be applied to discret-state random fields. Based on [42] this section provides a brief introduction to the algorithm-based and model-based hierarchical approaches, and each is categorized into different group of methods. Before describing each method, some notation and definitions for the hierarchical representation of random fields are provided in the next section.

2.5.1 Notation and definition

A random field Z is a graph $G = (S; \nu)$, where S is the set of nodes (the sites) and ν describe the relationship between the nodes (the neighborhood system). In the hierarchical representation of a problem a class of random fields Z^k are considered when $k = 1, \ldots, n$, where $k = 1$ and $k = n$ denote the coarsest and finest scales, respectively. In the hierarchical approach, the graph representation of a random field becomes $G = (S^1, \ldots, S^n; \nu)$. The neighborhood system in the hierarchical representation can include the spatial neighborhood system (the relationship between sites at each scale), or the inter-scale neighborhood system (the relationship

between sites at two or more different scales). Based on the neighborhood system, a tree-based or graph-based hierarchical framework can be defined. In the tree-based framework, scale Markovianity is considered. Therefore, each node is connected to the nodes at the next fine scale and no spatial neighborhood system is considered. Thus, the tree-based framework by definition does not include any cycle. However, in the graph-based framework in addition to the inter-scale relationship, spatial neighborhood system can be considered, therefore, scale and spatial Markovianity is observed. Fig. 2.7 shows examples of tree and graph-based hierarchical frameworks. The energy-based hierarchical models can be described by a tree or graph-based framework, in such models the inter-scale or spatial relationship can be described by the energy function in the Gibbs probability distribution. Depending on the nature of the problem, one can define separate energy models for each scale as H^k or choose a global energy model $\mathbf{H}$ for the whole hierarchy.

In the next two sections, four types of energy-based hierarchical approaches are described from the algorithmic and modeling perspective. Each approach is described briefly along with a schematic representation.

2.5.2 Algorithm-based hierarchical approaches

The algorithm-based hierarchical approaches are developed to simplify the optimization and synthesizing problems in image processing, such as restoration, segmentation, and sampling. These methods are based on two techniques: using an ordered relationship on sites by a stochastic or determinist algorithmic transformation on the set of sites, known as *renormalization group* [42, 44], and considering the problem as an ordered sequence of configuration spaces [45].

Renormalization group

The renormalization group method generates a multilevel structure by a multi- scale coarse-to-fine processing. Each scale is related to the other with a transformation such as decimation and block scaling (down-sampling). In decimation, a set of sites at a given scale is a subset of site at the previous finer scale. In block scaling, a site in a given coarse scale is a function (for example averaging) of a set of sites at the previous fine scale. Then given the measurement M, the posterior probability at scale k is defined as

$$Pr(Z^k = w^k|M) = \sum_{x^{k+1}} Pr(Z^k = w^k|Z^{k+1} = x^{k+1})\ Pr(Z^{k+1} = x^{k+1}|M), \quad (2.30)$$

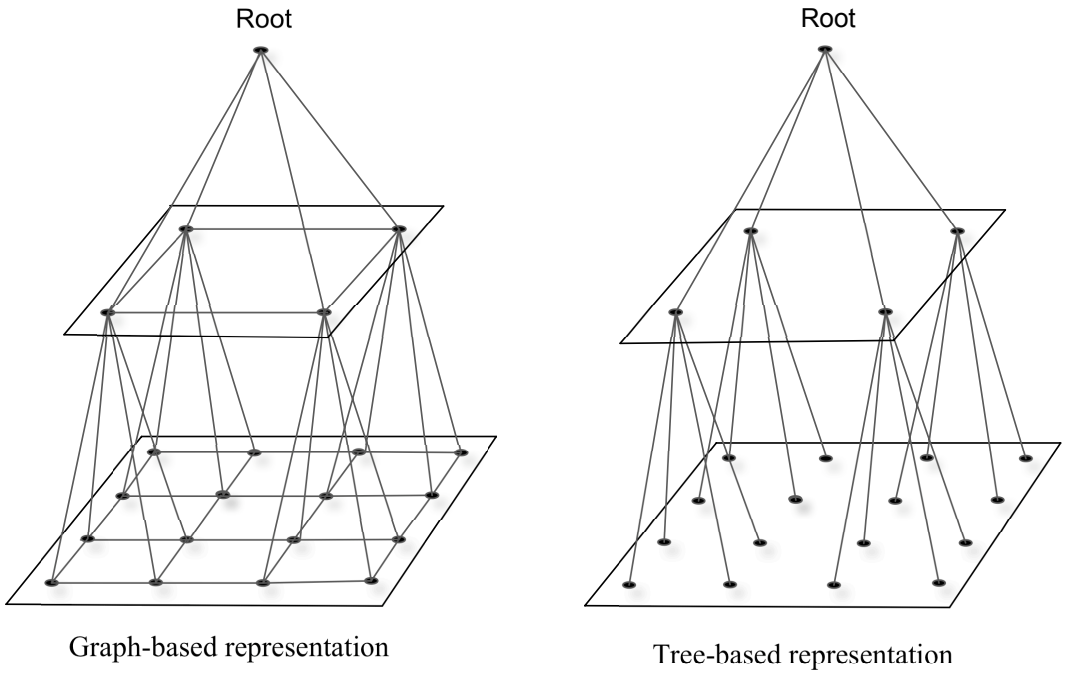


Figure 2.7: Graph-based representation versus a tree-based representation of the hierarchical Markov random fields.

where $w^k \in \Omega^k$ and $x^k \in \Omega^{k+1}$, and Ω^k and Ω^{k+1} are the configuration spaces at scale k and $k+1$,respectively. The probability function $Pr(Z^k = w^k | Z^{k+1} = x^{k+1})$ is defined based on a nonlinear transformation, such as decimation or block scaling.

Given the aforementioned nonlinear transformations, however, in general the probability $Pr(Z^k = w^k | M)$ is no more a local Gibbs probability for the coarse scales, i.e. it can not be written as the multiplication of the marginals, because the renormalization of the sites causes long range interaction to appear at the coarse scales [42, 44]. Therefore, for more convenience the probability is approximated with $Pr(Z^k = w^k | M^k)$ where M^k is an estimate of the measurement at scale k.

Based on the renormalization group approach, the energy is derived at each scale using the fine scale information. Then starting at the coarsest scale, a successive coarse-to-fine optimization approach is used in which the estimate at each scale is obtained by maximizing the posterior distribution, usually using simulated annealing. The estimate, $\hat{Z}^k$, is the solution of an energy minimization problem at scale k. Fig. 2.8 shows a schematic overview of the renormalization group approach.

Ordered-constrained configuration spaces

The ordered-constrained configuration spaces, named by Preteux et al. in [41, 42, 45] is motivated from multi-grid techniques [43], and is based on solving the global optimization problem by a sequence of subspaces of the original space of configuration. Staring with the coarsest scale, at each scale an estimate is obtained using optimization algorithms such as simulated annealing, then the estimate is mapped into the next scale and considered as an initialization guess for the optimization problem at the finer scale. This procedure continues until the finest scale is reached. Despite the renormalization group approach, the ordered-constrained configuration approach preserves the Markovianity of the random fields at each scale, because at each scale a separate model is defined, independent from the fine scale, and also each scale is constrained by the measurement M. Therefore, the estimate $\hat{Z}^k$ at scale k is obtained as

$$\hat{Z}^k = arg \min_{Z^k} H^k(Z^k; M). \tag{2.31}$$

and the estimate is mapped into the next scale using an interpolation function $D(\cdot)$. Fig. 2.9 shows a schematic representation of the approach based on Ordered-constrained configuration spaces.

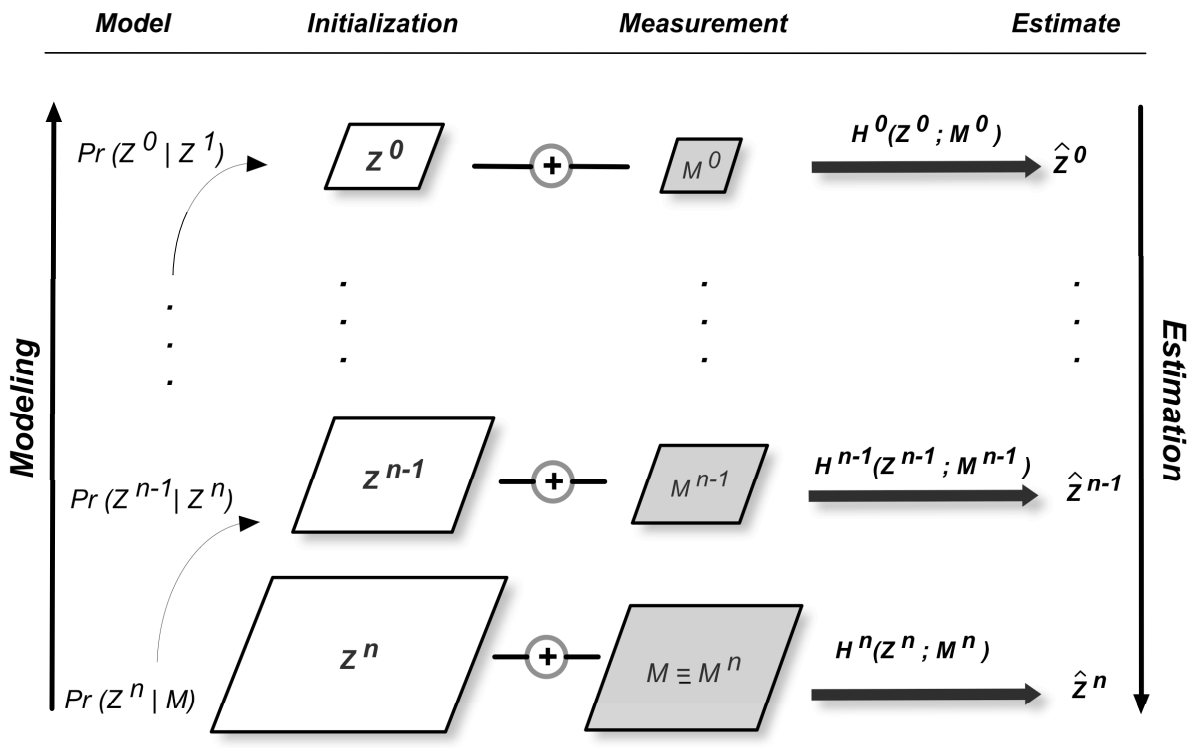


Figure 2.8: Renormalization group hierarchical method. The energy model at each scale is learned through a bottom-up approach and based on the conditional distribution between random fields at two consequent scales and the measurement is estimated at each scale.

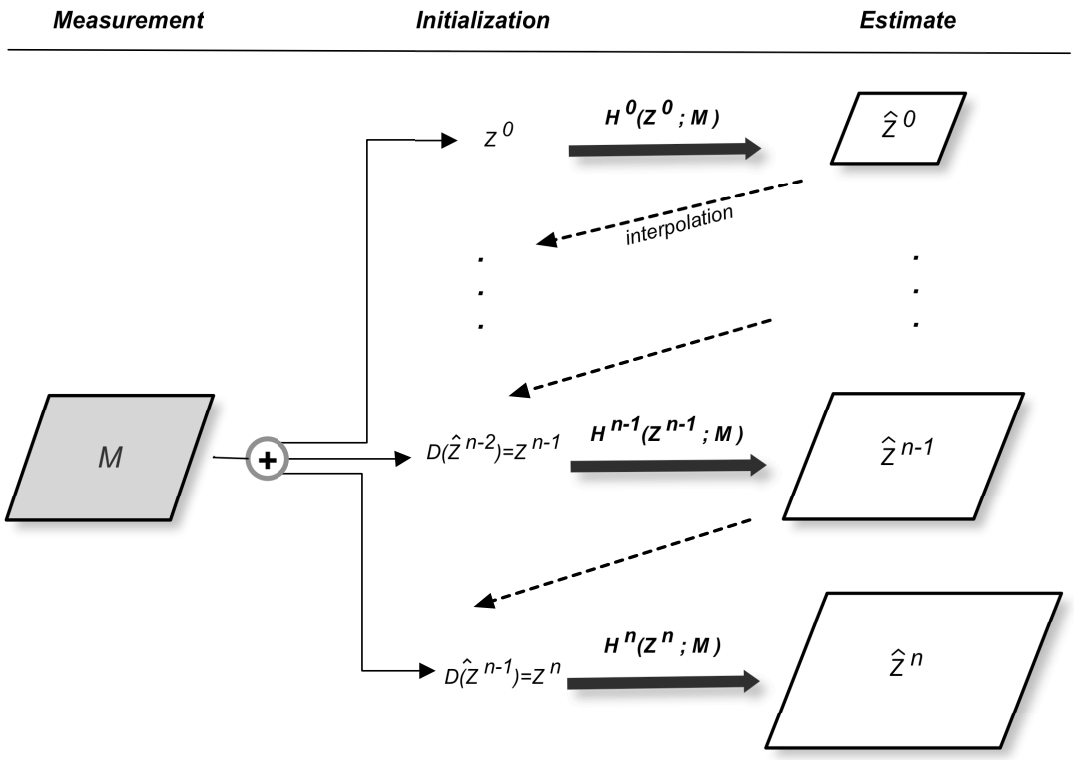


Figure 2.9: The hierarchical approach based on ordered-constrained configuration spaces. A separate energy function is defined for each scale. The estimate at each scale is obtained by optimization algorithms and used in the fine scale as an initialization guess.

2.5.3 Model-based hierarchical approaches

Based on the model-based hierarchical approach, a new global model or an ordered family of models are defined for the whole hierarchical graph. In model-based approach a global Markov model is considered for the whole hierarchy, therefore the scale Markovianity is inherent in the model, while in the algorithm-based hierarchical approaches the Markovianity is considered at each scale rather than between different scales. The hierarchical models based on this approach use a graph or tree-based representation as shown in Fig. 2.7. There are two main categories for the model-based hierarchical models based on using a single Markov model or a family of models.

Hierarchical graph-based models

Here, a random Markov process is defined for the whole hierarchy and global energy is associated with that. The global energy can be defined either by the sum of local potentials at each scale, leading to spatial Markovianity, and/or by a transition probability from one scale to the other, which interrelates the sites at each scale with the sites at the finer or coarser scale. Therefore, the prior distribution is

$$p(Z = w) = p(Z^0 = w^0) \prod_{k=1}^{n} p(Z^k = w^k | Z^{k-1} = w^{k-1}) \tag{2.32}$$

Fig. 2.10 shows a schematic representation of the hierarchical graph-based models. Viterbi-like algorithm, or Kalman filer-like algorithms can be used to solve the optimization problem in this case.

Ordered family of models

In the ordered family of models a family of processes are considered, each associated with a single scale in the hierarchy. Given the original Markov model of energy $H(Z; M)$, a similar Markov model is defined at each scale as $H^k(Z^k; M^k)$, where M^k is obtained using multi-resolution techniques such as wavelet or Gaussian filter decomposition. The estimation problem is solved using a coarse-to-fine exploration of the hierarchical structure, by starting from the coarsest level. Using energy minimization methods such as simulated annealing, an estimate $\hat{Z}^k$ at a coarse scale is obtained by minimizing the energy model $H^k(Z^k; M^k)$ and then the estimate is propagated into the next scale. In most of the cases the estimate is interpolated at the next scale, and it can be used as an initialization guess for the energy

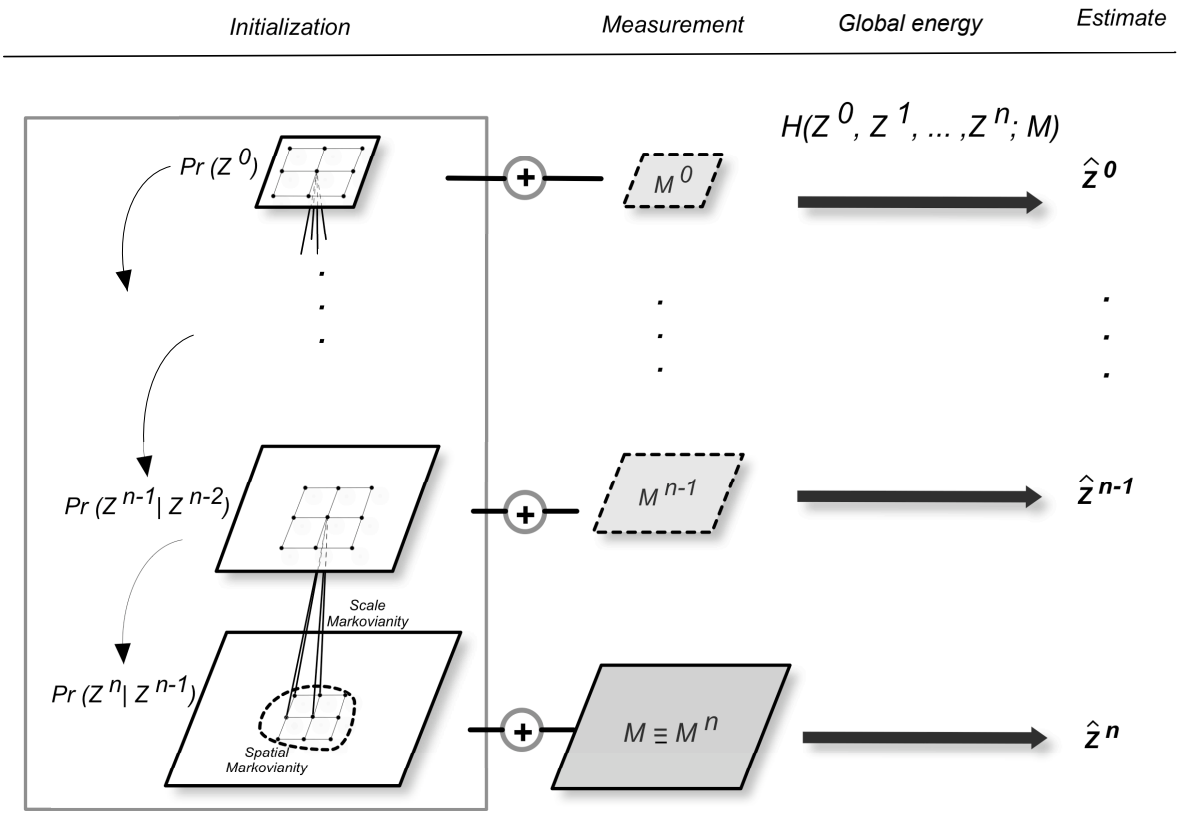


Figure 2.10: Graph-based hierarchical approaches. A global energy H is defined for the whole framework and then by estimating the measurement at each scale, the global estimation problem is solved using Kalman filer or Viterbi-like algorithms.

minimization problem at that scale. Another approach in propagating the estimate $\hat{Z}^k$ at scale k into scale $k+1$ is to add $\hat{Z}^k$ with the estimate obtained at scale $k+1$. At the end, if $\hat{Z}$ denotes the ultimate estimate of the problem, then

$$\hat{Z} = \hat{Z}^n + \sum_{k=0}^{k=n} D(\hat{Z}^k) \tag{2.33}$$

where $D(\cdot)$ is an interpolation operator, and n is the number of scales. Fig. 2.11 shows a schematic representation of the ordered family of models hierarchical approach. Although, the model at each scale is the same as the global Markov model, there are other situations in which different Markov / non-Markov models is considered in such hierarchical framework [42].

For the purpose of this research, we have considered a model-based hierarchical approach which is similar to the ordered family of models framework discussed here. Chapter 6 describes the hierarchical framework that has been used in this research in details.

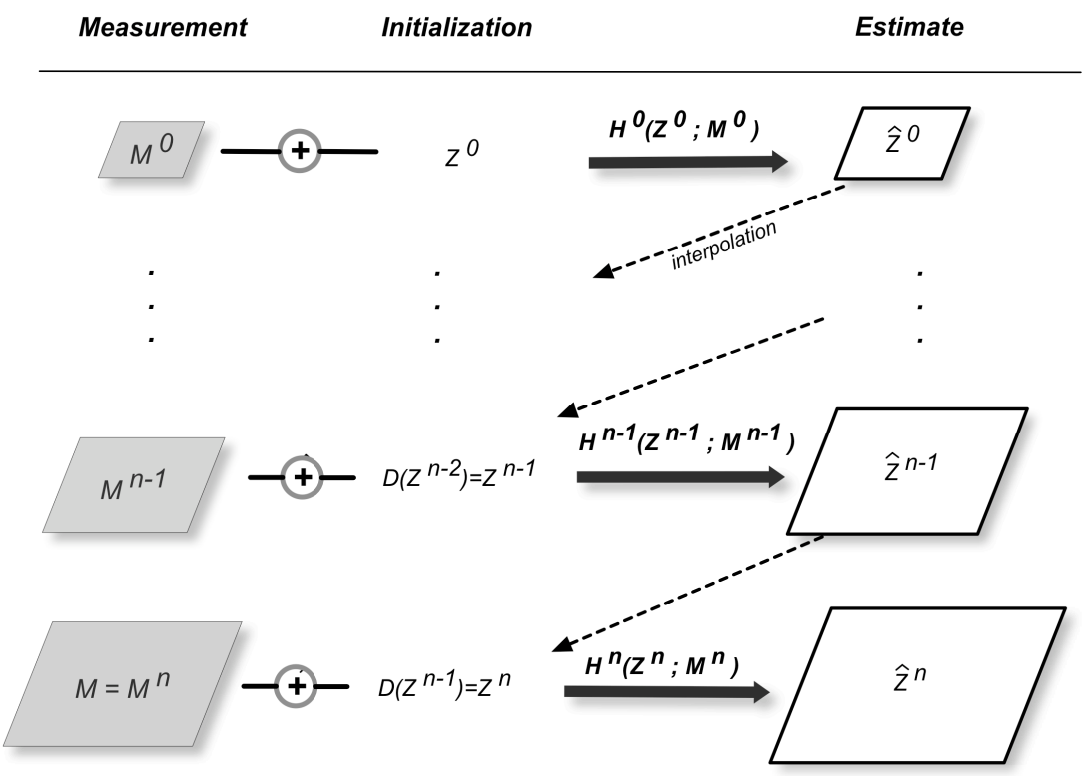


Figure 2.11: Hierarchical approach based on ordered family of models. A separate Markov model is considered at each scale and the estimate at each scale is feed into the next scale using interpolation or incremental processes.

2.6 Summary

This chapter has provided a brief overview of the background materials required to develop the contributions in this thesis. The chapter starts with brief introduction of sampling and estimation problems and continues with a definition and overview of porous media reconstruction. Then, different modeling approaches that can be used to model the complex structures of porous media are discussed. The MCMC sampling methods along with simulated annealing is reviewed. Finally, energy-based hierarchical modeling approaches for random fields representations are discussed briefly.

Chapter 3

Problem Formulation

In this chapter the main research idea in the thesis is presented. The chapter starts with a literature review on porous media reconstruction to motivate the reader, then continues with defining, describing and formulating the problem. Then, different aspects of the proposed approach to solve the problem are discussed and finally, in the last section some pre-liminary results are provided.

3.1 Porous Media Reconstruction

Problem Statement: Generating artificial samples of porous media images is a crucial process in studying different properties of porous materials. The artificial samples are generated from a statistical model learned from a limited set of high resolution training data, using a porous media stochastic reconstruction process. The stochastic reconstruction is a sampling process by which *typical* artificial samples of porous materials are computationally generated. Depending on the absence or presence of measurements, the process is defined as *prior* or *posterior* sampling, respectively [6, 7]. The overwhelming majority of porous media simulation methods in the literature perform prior sampling, in which the reconstruction is based solely on the learned prior model [3, 6, 8, 9, 10, 11]. The prior model, as the ones described in Section 2.3, typically consists of one or more statistical functions (e.g. two-point correlation, chord-length distribution, etc.) learned from 2D images of the pore and solid space. However the growing availability and use of tomographic and MRI measurements (shown in Fig. 2.2), means that in many cases it is desirable to study a particular physical sample, in which case measurements are also available, and the reconstruction should reflect the pore and solid structures appearing

in the measured sample, thus a *posterior sampling* approach [12] is needed. This approach is based on *fusing* the information presented in the measurements and the prior model.

Fusing low resolution measurements with high resolution data has been studied recently in porous media literature. Considering the measured porosity as an initial condition, Zhao et al. [25] have proposed a stochastic reconstruction method based on combining two-point correlations and lineal path functions. Okabe and Blunt [5] have included low resolution information obtained from tomographic images of porous media in which only large pores are resolved. In their method, the large pores are reconstructed based on the tomographic data while the unresolved small scale structures are generated according to the prior model, *independently* from the large pores. There are two problems with this approach

- the measurements are not explicitly coupled with the prior model in computing a reconstruction sample,
- only a small portion of the information provided by the low resolution 3D measurements is exploited, namely the overall sample porosity and the presence of pore and solid domains.

More specifically, low-resolution measurements, can be rich sources of information. For instance, X-ray computed tomography or MRI measurements resolve not only the relatively larger pores, but also provide information on the local (voxel-scale) porosity of unresolved structures at different scales. In addition, MRI can provide 3D maps of parameters sensitive to the geometry and connectivity of unresolved scales [46]. For example, Pomerantz et al. [47] have recently presented spatially-resolved measurements of the decay of transverse magnetization in a porous material and analyzed them to obtain the spectrum of decay constants, T_2, at a voxel. The physical interpretation of a T_2 spectrum as a distribution of pore surface-to-volume ratio, S/V, provides additional information on the geometry (pore size) of unresolved structures [4, 48]. Therefore, in addition to the local porosity information provided by the MRI measurements, the measurements of the pore size at different scales can enrich the stochastic reconstruction process. Both types of measurements can be fused with the prior model to reflect different statistical features and aspects of a given material.

As motivated by the stochastic reconstruction of porous materials, the challenge here is how to fuse the prior model with different types of measurement or constraints, to form a posterior model and then how to generate samples from the posterior model.

3.2 Modeling

Image reconstruction, which is an inverse problem, can be viewed as an image resolution enhancement problem in which multiple observations are fused to generate a data with better resolution and possibly less noise. Based on the presence or absence of prior knowledge, the reconstruction task is done using a Bayesian or non-Bayesian approach, as described in section 2.1. In a non-Bayesian case (Fig. 3.1 (a,b)), multiple observations can be combined using methods of super resolution and image fusion [49] to generate high-resolution realizations. The super resolution methods are based on combining multiple measured images taken from a same scene, to generate an image with higher resolution from the measurements. The multiple measurements are the same except for a sub-pixel shift, rotation or noise [49]. However, in the case that the luxury of multiple images are not available, such as in porous media reconstruction, this approach is not applicable. In addition, the super resolution methods are based on obtaining an estimate using estimation methods described in Section 2.1.1. While in porous media reconstruction we are facing a sampling problem, seeking a typical realization drawn from the prior or posterior model [7].

In multi-resolution image fusion approaches [16] (Fig. 3.1 (b)), the image data are decomposed into multiple scales using multi-resolution decomposition methods such as wavelet transform. Then the image data at multiple resolutions are fused using pixel-based or region-based fusion algorithms [16], and finally the fused image is transformed into the original domain. The multi-resolution image fusion approaches assume that the images are obtained from the same scene, but with different modalities – an assumption which is not true in porous media reconstruction. The porous media images which are going to be fused are obtained from the same material rather than the same scene, so they need to obey the same statistical features. Therefore, the multi-resolution image fusion techniques are not practical here, rather we need to define a statistical fusion approach to combine the information obtained from the high resolution and low resolution images of porous media, as shown in Fig. 3.1 (c).

The limited number of 2D high resolution samples of a porous material contain valuable and critical information about the structures and distributions of pores and solids, although they are limited in number and they can not be applied directly for image enhancement task. The high resolution samples are used as training data to define statistical descriptors and to learn the parameters of the models (such as histogram or chord-lengh) which are usually used to describe the statistical features and characteristics of such materials. The reconstruction task uses these models as

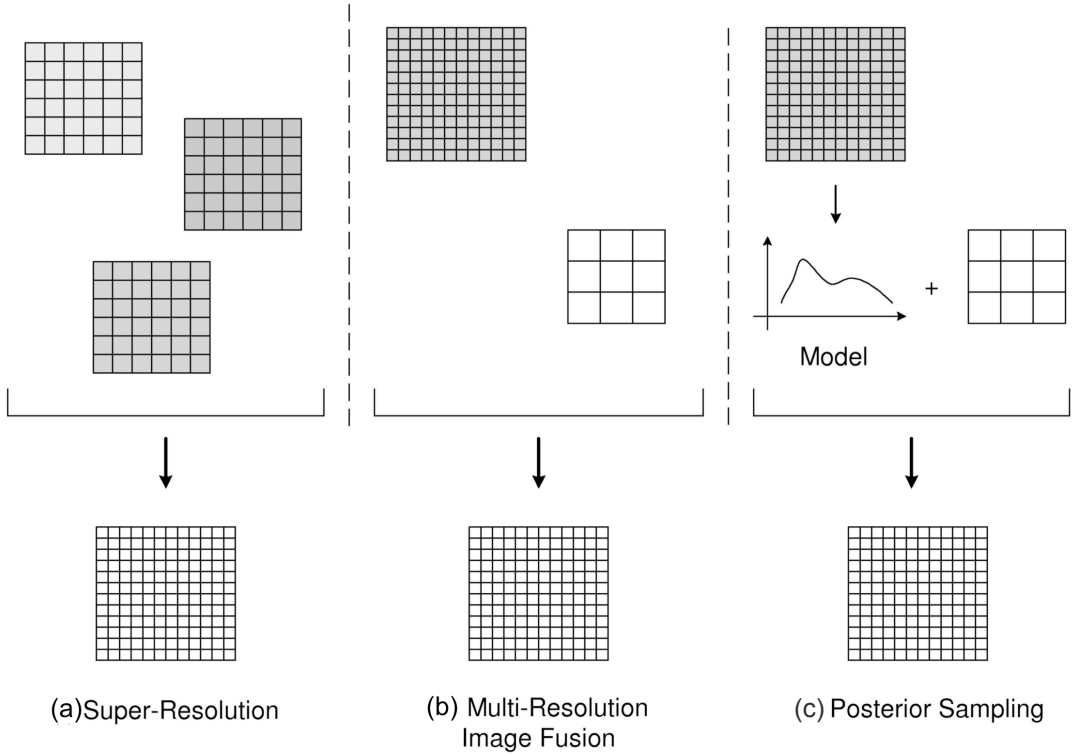
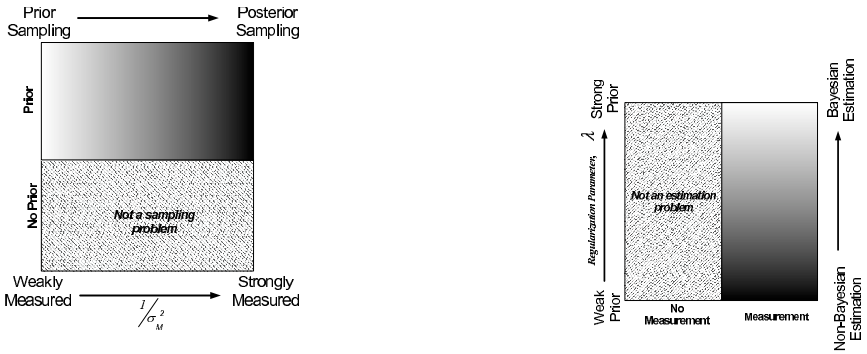


Figure 3.1: Different approaches in data fusion: We may be registering and fusing multiple images (a), fusing data across resolution (b), or statistical fusion through a model (c), as we propose in this paper. Although (b) and (c) appear superficially similar, in (b) the two data sets correspond to the identical underlying images, whereas in (c) the data sets obey the same statistics

a priori knowledge to generate artificial samples.

Considering the prior model obtained from the 2D high resolution samples, we propose a Bayesian image reconstruction approach by which the low resolution measurements such as MR images are coupled with the prior model to generate artificial samples (Fig. 3.1 (c)). In this approach there is no need to have the measurements and high-resolution realizations from the same original image, since rather than fusing the images physically, we consider that the images share the same statistics and consequently we fuse the images *statistically*.

The measurements available for reconstruction may be (i) weakly constrain-



(a) Sampling (b) Estimation

Figure 3.2: Sampling and estimation. In sampling (a) there is a continuum between prior and posterior sampling, depending on the measurement noise variance σ_M^2 – the degree to which the unknown random field is constrained by the measurements relative to the prior. In estimation (b) there is a continuum between Bayesian and non-Bayesian problems, depending on the regularization parameter λ – the degree to which a prior model is present to constrain the random field.

ing, (ii) partly constraining, or (iii) fully determining, a continuum illustrated in Fig. 3.2(a). A great many problems in image processing fall into (iii), such that the measurements are of such a resolution and quality that no prior is needed, and the image is processed or reconstructed in the absence of any specific prior. At the other extreme (i) the measurements are so weak that one is essentially randomly sampling from a prior model.

In principle, given measurements and a prior model, reconstruction can proceed as either a sampling or estimation problem, as illustrated in Figure 3.2. Since our measurements fail to resolve many of the fine scale details and at the same time fine structure must be present in the stochastic reconstruction, a sampling approach is required [7].

Fig. 3.3 shows how the posterior sampling approach combines the measurements with the high resolution samples to generate reconstructed porous media samples. As the unknown, training sample and themeasurements are assumed to be from the same material they have the same statistical features.

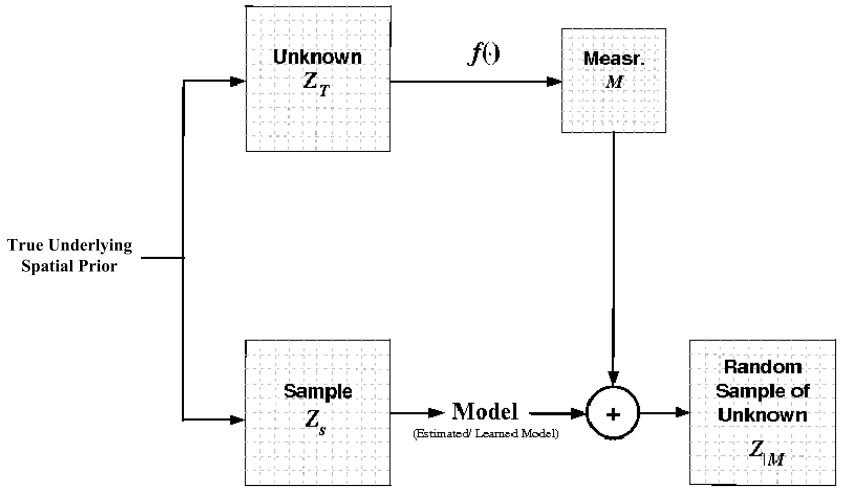


Figure 3.3: Posterior Sampling: The sample image Z_s is used to learn the prior model. A low resolution measurement M is taken from the unknown image Z_T. The unknown image Z_T and the sample image Z_s obey the same statistics.

For the statistical fusion problem, we are left with defining

- *the prior model, $p(\cdot)$, which reflects statistical features and characteristics of the image. The prior model will be defined using the 2D high-resolution samples,*
- *the constraint term, M, usually contains measurement which is an incomplete/irregular observation of the original data.*

Then the problem is how to generate samples which are statistically consistent with the prior model and the measurements.

Therefore, given M we want to draw sample Z from the posterior probability distribution, $p(Z|M)$,

$$p(Z|M) \propto p(Z)\ p(M|Z) \tag{3.1}$$

While simple stochastic image models such as correlation models and spatial variance [27] can be used to model porous media, they are very poor due to the

chaotic and complex morphology of these material, and instead for discrete-state problems (porous media images are binary) widely-used GRF are considered. Using Gibbs random fields theory, the prior model can be described by the Gibbs probability distribution, $p(\cdot)$. Combining the measurement with the prior leads to the posterior probability distribution, as shown below

$$\begin{array}{ll} \textit{Prior} & \textit{Posterior} \\ H(Z) = \mathcal{H}(Z) & H(Z|M) = H(Z) + \alpha J(Z;M) \\ p(Z) = \dfrac{1}{\mathcal{Z}}\, e^{-\frac{1}{T}H(Z)} & p(Z|M) = \dfrac{1}{\mathcal{Z}_M}\, e^{-\frac{1}{T}H(Z|M)} \end{array} \tag{3.2}$$

where $J(\cdot)$ is the constraint term, describing how the measurement M is incorporated with the model and α is a parameter balancing the the influence of the measurement and prior in the energy function.

The way that the prior and posterior energy function are defined, is discussed in the next two sections.

3.2.1 Prior model

The prior model can be studied in terms of three aspects:

- accuracy: determines whether the prior model reflects all the characteristics of the original data perfectly,
- complexity: determines whether the prior model is mathematically complicated or complex,
- existence: determines whether the inherent nature of the original data obey a prior model, i.e. whether there exists a prior model, or the inherent nature of the problem does not obey a prior model and an artificial prior model should be considered for the purpose of regularization.

Porous media have inherent characteristics and features that can be asserted by a prior model citeAdler,Torquato,Sobczyk. However, due to the complex nature of porous materials, the mathematical prior models used for describing their behavior are not able to characterize all of their properties. By combining different prior models, one is able to characterize more complex structures than a single model. For instance, one can combine histogram model with the chord-length model to

capture the local small structures while describing the length scale of pore and solid space [25, 50].

The Gibbs prior probability distribution is fully characterized by the prior energy $H(\cdot)$. There are various possibilities for H, such as histogram, chord-length and/or correlation models described in Section 2.3. Depending on the nature of the problem, we can merge multiple prior models to reflect various aspects of the data. Chordlength and histogram models are widely used in porous media reconstruction, here we define the prior energy for these two models. It is obvious that the energy function can be defined for other types of models as well.

Based on the histogram model defined in Section 2.3.3, the energy function when using the histogram model is defined as

$$H_h(Z) = \sum_{k=1}^{2^9} \left(\frac{\|\bar{h}(k) - h(k, Z)\|}{v(k) + \epsilon} \right) \tag{3.3}$$

where $\bar{h}$ is the learned histogram distribution, and $h(k, Z)$ is the observed distribution for a given, simulated Z. The term v is the variance for the respective histogram entries to account for sample variation. A small constant ϵ is introduced to avoid divisions by zero, especially in the comparatively common case of unobserved configurations k corresponding to $\bar{h}(k) = 0$.

The chord-length distribution function $C^i(\ell)$ [3], as defined in Section 2.3.4, can be interpreted as the probability of finding a chord with length ℓ in phase i. For a two-phase porous medium, chord-length distributions can be defined for either phase and for chords at different orientations. It is common to limit the orientation to the horizontal and vertical directions in 2D, in which case the chord length energy in phase i is

$$H_c^i(Z) = \sum_{k=0}^{L} [\|\bar{C_h^i}(k) - C_h(k, Z)\| + \|\bar{C_v^i}(k) - C_v(k, Z)\|] \tag{3.4}$$

where $\bar{C_h^i}$ and $\bar{C_v^i}$ are the learned model for the horizontal and vertical chords, respectively, and C_h^i and C_v^i are the chord-length distributions of a simulated random field Z. The norm function $\|\cdot\|$ is chosen to be an ℓ^2 norm which has a reasonable computational complexity.

3.2.2 Measurement model

According to Eq. (3.2), the constraint term J in the posterior energy function defines how measurements are formulated and asserted in the model. In order to

determine J we need to

- define the forward problem function, f, which according to Section 2.1, determines how measurements are related to the original data,
- define a relationship in J to include M and f.

The forward model, f, differs from one problem to another and depends on how the measurements are generated. For instant, the low resolution measurements of MRI can be interpreted as the down-sampled version of the original data. By down-sampling we mean that a given datum, $M_{i,j}$, in the measurement space represents the average of a corresponding block of pixels, $B_{i,j}$, in the original data space, therefore

$$M_{i,j} = f(\{Z\}_{i,j}) = \frac{1}{d^2} \sum_{h=i*d+1}^{i*d+d} \sum_{g=j*d+1}^{j*d+d} Z_{h,g} \tag{3.5}$$

$Z_{h,g}$ denote a pixel in the original image, and d is the down-sampling parameter, such that for an $n \times n$ original image and an $m \times m$ measurement, then $d = n/m$. The set $\{Z\}_{i,j}$ is the set of pixels in the original data space, corresponding to the pixel $M_{i,j}$ in the measurement space.

Then for a given sample Z, we can incorporate the measurement and the above forward model in J, as

$$J(Z; M) = \|f(Z) - M\| \tag{3.6}$$

where $\|\cdot\|$ is the ℓ^2 norm which has a reasonable computational complexity, and it determines how much a given sample Z is consistent with measurement. Assuming that the noise in the measurement is Gaussian, the choice of ℓ^2 norm in the constraint is reasonable, however depending on the nature of the problem and the statistics of the measurement, one can choose other types of norm functions as well.

Depending on the nature of the problem and how we are going to involve the measurement in the model, different forms for J can be defined, for example in the hierarchical formulation of the problem that we are going to discuss in Chapter 6, other forms are defined. In this section the statistical model for fusing the prior model with the measurement has been described. In the next section we discuss how we can generate posterior samples from the proposed posterior model.

3.3 Posterior Sampling and Annealing with Constraints

After defining the prior model and the constraint term J in Section 3.2, here we describe how to choose the parameters T and α in the posterior model to generate random samples. As shown in Eq. (2.6), posterior sampling produces results similar to the solution of estimation problem (Eq. (2.5)) in the area that the measurements are dense, and generates a random synthesis in those areas not constrained by measured values. In fact, the measurements in the posterior distribution, constraint degrees of freedom of the prior model.

3.3.1 Sampling from Gibbs probability distribution

In principle, to generate random samples, the posterior sampling appears straightforward: specify the energy function and run the Gibbs sampler as described in Section 2.4.

In practice the problem is not at all straightforward. Astonishingly, almost all porous media MCMC papers [31, 51] do simulated annealing to generate "random samples" from the prior model, a procedure which succeeds because the annealing process fails to find the optimum Z maximizing $p(Z)$, and finds a random, near-optimum Z instead. But here lies the problem:

- We are not seeking the most probable realization, rather we want a *random* realization faithful to the prior model and measurements, that is, we wish to draw a random sample from $p(Z)$ at $T \neq 0$.
- However the energy function $H(Z)$, empirically derived from porous media, is really only valid at $T = 0$. That is, Z is unlikely to look like a porous medium unless the empirical *constraints* in the energy function are asserted.

The theory of annealing and sampling with constraints has been reviewed before in the literature by Geman [29]. Based on this theory, sampling from the Gibbs probability distribution can proceed at a non-zero temperature $T \neq 0$ while increasing the parameter α slowly enough [29] to infinity. On the other hand, decreasing T and increasing α simultaneously, yield the solution of an optimization problem in which the most probable realization is obtained. More specifically,

- If T is constant, $\alpha \to \infty$, and $\lim \frac{\alpha}{T} = \epsilon$ is sufficiently small, then we are generating a sample $Z_{|M}$ from the constrained space, such that

$$Z_{|M} \in \{Z \mid J(Z; M) = 0\} \tag{3.7}$$

  In this case the constraint term J is rigidly asserted.

- If $T \to 0$, $\alpha \to \infty$, and $\lim \frac{\alpha}{T} = \epsilon$ is sufficiently small, then we get the estimate $\hat{Z}$ minimizing the posterior energy function, such that

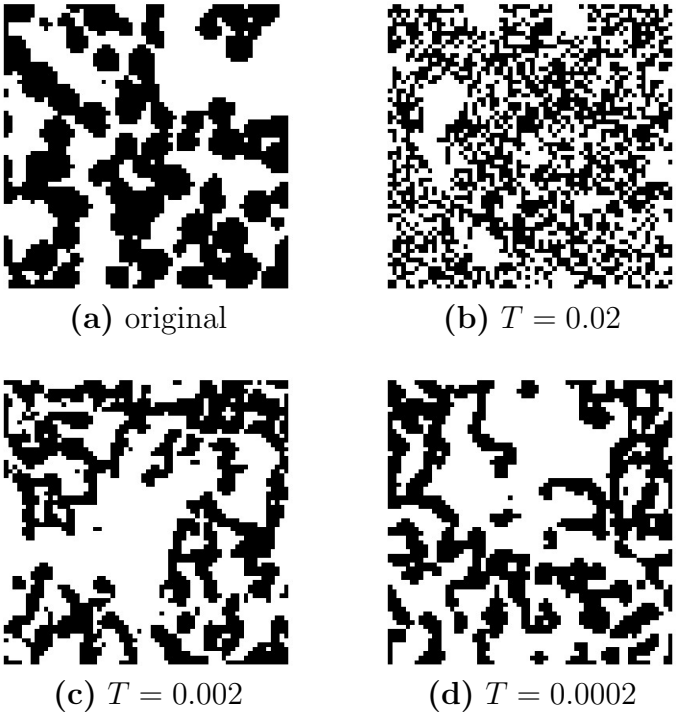
$$\hat{Z} = arg \min_{Z}(H(Z) + \alpha J(Z; M)) \tag{3.8}$$

In practice, for porous media reconstruction, the temperature T can not be considered as a fixed value at the beginning, since we do not know at what value for T we should run the sampler to generate random samples. This uncertainty stems from how energy models $H(Z|M)$ are developed for Z — we really do not know any "true" model, rather we infer plausible constraints from sample data, such that implausible $Z_{|M}$ is associated with a larger energy H, and plausible $Z_{|M}$ with a small value of H. Because of the highly qualitative notions of plausibility and implausibility, it is hardly possible to select a value of T which leads to an accurate posterior distribution $p(Z|M)$. This fact can be seen in Fig. 3.4. Samples in Fig. 3.4 are obtained by doing incomplete annealing, i.e. decreasing T to a specific point in the prior energy function. Comparing the case for $T = 0.02$ with the case for $T = 0.0002$, the former case does not possess porous media structures, and the later case entails less noise but still does not have the same pore structures as the original image.

The posterior sampling problem is most effectively understood in the context of simulated annealing [29]. As discussed in Section 2.4.3, simulated annealing runs a Gibbs sampler, initially at a high temperature, where the image contains purely random structures, then at progressively lower temperatures until the system reaches a converged state.

Simulated annealing algorithm guarantees the convergence to the optimal solution *if* the temperature is decreased in the logarithmic form defined in Eq. (2.28). However, as described in Section 2.4.3, the logarithmic schedule is not practical, and usually a faster annealing schedule is used such as exponential schedule, defined in Eq. (2.29). Using any other faster schedule than the logarithmic schedule leads to *quenching* (cooling rapidly) rather than annealing. Therefore, some of the invalid small-scale structures in Fig. 3.4 are due to rapid cooling rather than annealing.

Annealing over the posterior energy, as proposed in this research, is also based on using an exponential annealing schedule, and consequently can be viewed as a



(a) original **(b)** $T = 0.02$

(c) $T = 0.002$ **(d)** $T = 0.0002$

Figure 3.4: Samples generated by finite annealing. (b)-(d)Different samples generated from the prior model, by annealing down to $T > 0$.

quenching approach. However the constraint in the energy function is asserting additional information which avoids the sample to encompass invalid structures which might be generated due to quenching. This problem will be re-visited in more details and from other perspectives in Chapter 5.

3.3.2 Choosing the parameters T and α

Inspired by the annealing with constraint approach presented in [29] and in Eq. (3.7), we propose to start with a high temperature and decrease the temperature slowly according to the exponential schedule, since as described before we can not come up with a fixed value for T. However, due to possible noise and uncertainty in the measurement, we do not want to assert the constraint rigidly, but rather we are facing with a *soft* constraint situation. Therefore, we do not increase α, rather we consider a fixed value for α, which is equivalent to a very slowly increasing rate as compared to decreasing T.

The parameter α is chosen based on the initial values of the constraint and the prior, such that

$$\alpha = c\,\frac{H_0}{J_0} \tag{3.9}$$

where H_0 and J_0 are the initial energy values for the prior and measurement energy, respectively, and c is a parameter balancing the weight of the measurement in the posterior energy function. There are three possible approaches to setting c

1. An initial balance by setting $c = 1$: the prior and constraint energy are having the same contribution at the beginning of the annealing process
2. An initial dominance by the prior model by setting $c < 1$: the prior energy has more contribution than the constraints term
3. An initial dominance by the constraint by setting $c > 1$: the constraint term has more contribution than the prior term.

Therefore, the parameter α can be set depending on the degree of belief in the prior and/or the measurement.

In Fig. 3.5 the prior energy during the annealing process for different choice of c is illustrated. When $c < 1$, the prior has more weight in the energy function, therefore it is going to be more effective, and we can see that the prior energy reaches a smaller value at the end, compared with the other cases. When $c > 1$, the measurement exerts a greater influence in the process and therefore the prior may not have enough chance to get as small as for the case with $c < 1$.

3.4 Preliminary Results

As a preliminary experiment, we have applied the proposed posterior sampling approach to a small set of porous media images with two different models: histogram and Ising. The results shown in Fig. 3.6 are generated based on an equivalent contribution from the prior and measurement, i.e. $\alpha = \frac{H_0}{J_0}$. The measurement model is that considered in Eq. (3.6) and the forward model is the same as the one described in Eq. (3.5), while the decimation factor d is equal to 8. As can be seen from the figure, the Ising model can not capture the complex structures of porous media, while the histogram model is more successful than the Ising in reconstructing the details.

The results in Fig. 3.6 are preliminary results. More detailed results and evaluations are presented in Chapter 4, 5 and 6.

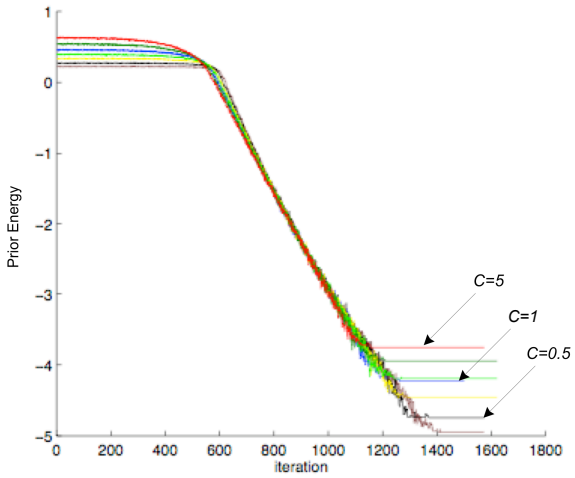
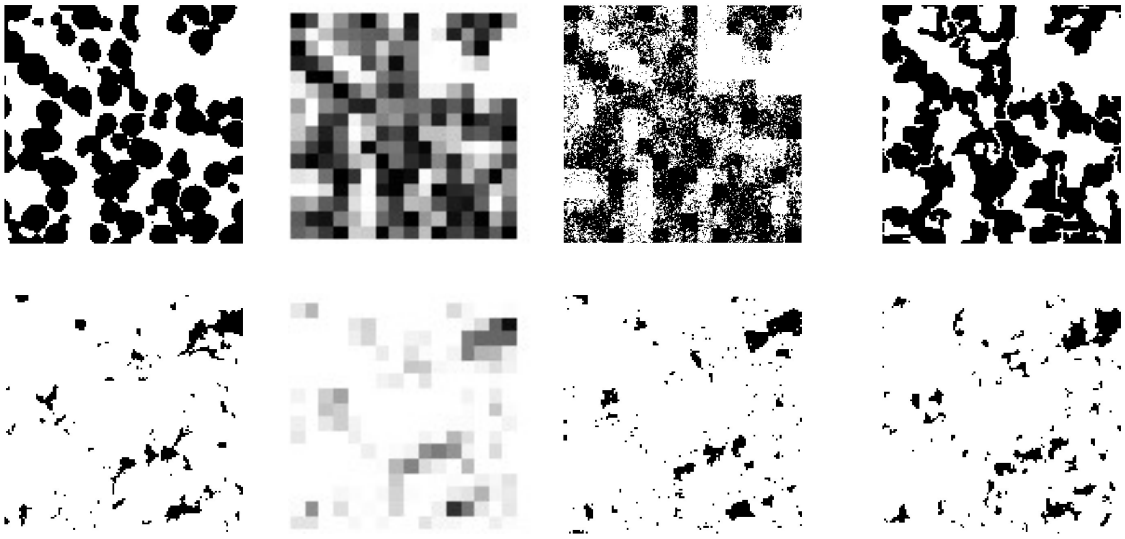


Figure 3.5: Prior energy as a function of iteration for different values of c in Eq. (3.3.2). As c gets larger than 1, the measurement term exerts a greater influence than the prior term in the process, and consequently does not allow the prior energy to decrease as for the case with $c < 1$.

3.5 Summary

In this chapter, a model-based framework for image sampling and reconstruction of porous media is introduced [12]. In contrast to the current approaches such as super-resolution and multi-resolution image fusion, we have proposed a Bayesian statistical framework in which low-resolution measurement constrains the prior model. We have introduced a posterior sampling approach using Gibbs sampler and simulate annealing for porous media reconstruction. Different strategies for choosing the parameters in the posterior model are discussed. At the end the reconstruction of low resolution porous media images are provided based on Ising and histogram prior model. The framework proposed in this chapter is extended in Chapter 4, 5 and 6 to handle more complex situations.



(a) Original images (b) Measurement (c) Reconstruction using Ising model (d) Reconstruction using histogram model

Figure 3.6: Reconstructed original image (a), given measurement as (b) and using two different models. The result using Ising and Histogram model is illustrated respectively in (c) and (d). The resolution of original image is 8 times greater than the measurement, i.e., $d = 8$.

Chapter 4

Posterior Sampling for Two-Scale Porous Media Reconstruction

As proposed in Chapter 3 the low resolution measurement usually obtained from MR imaging is combined with a prior model to form a posterior model. Since porous media are nearly fractal-like, with pores on all scales, low resolution measurements will fail to resolve the pores at fine scales. An interesting case occurs when only the large pores and/or solid structures are resolved by the measurements, such as the MRI measurements shown in Fig. 2.2. Motivated by this special case, in this chapter a data fusion approach is proposed in which different types of low resolution measurements are coupled with a prior model for reconstructing two-scale porous media images.

4.1 Problem Statement

As discussed in Section 3.2 and shown in Fig. 3.2, based on how much information (the resolution) the measurements are providing, the porous media reconstruction fall into a continuum from estimation to prior sampling. Our interest lies in between, such that something of a delicate balance is required between the assertions of measurements and prior model. This is the case with MRI or X-ray computed tomographic data of porous media, in which relatively large pores are explicitly resolved but smaller scale structures appear in the measured images as different shades of gray. Although the low resolution measurements such as MRI data only resolve the large pore, such measurements are able to provide valuable clues on the

local porosity (fraction of pore space) for the unresolved structures. The reconstruction of porous media in this case is a two-scale reconstruction process, since at least the pores at one of the scales (the largest scale pores) are resolved.

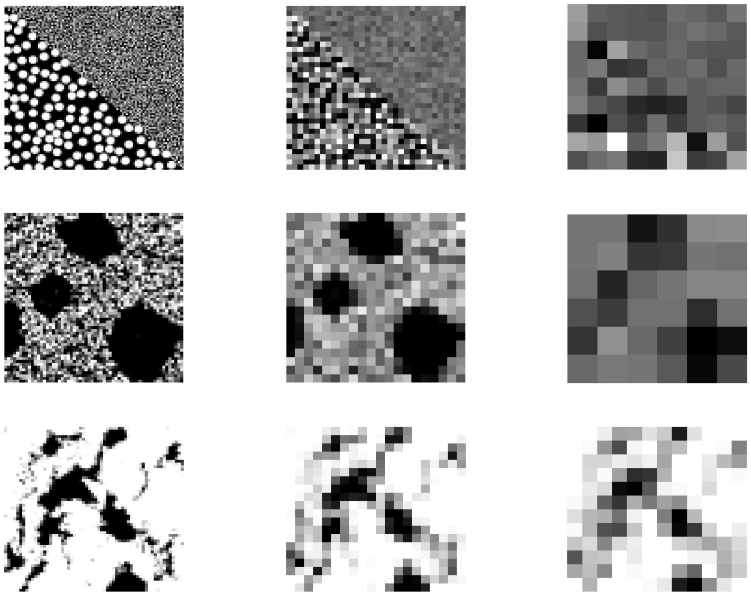
> *The challenge in a two-scale porous media reconstruction is how to couple the measurements with the prior and possibly incorporate other types of measurement which can provide information on the geometry of pores/solids to enrich the reconstruction process.*

To cope with this challenge, in the next two sections we define a measurement model and combine it with the prior model to form a posterior model and then introduce a sampling approach based on the approach discussed in Chapter 3, to sample from the posterior model.

4.2 Measurements as the Constraints

A typical low resolution measurement of porous media obtained by computed tomography or MRI measures the *local* porosity (the fraction of void) of a given medium. The low resolution measurement, M_p, measures the porosity of a given material within a pixel in 2D or a voxel in 3D. Therefore, each datum in the porosity measurement corresponds to the pore fraction of a given block of pixels of the underlying high resolution field Z.

Fig. 4.1 plots porosity measurements at three different resolutions for three fields, the first synthetic, the second from sintered glass beads and the third from carbonate rock. The down-sampling parameter d measures the reduction in resolution from Z to M, such that each single measurement measures $d \times d$ elements in two-dimensional Z. In (a) the measurements resolve all aspects of the field, leaving little to do for a prior model, whereas in (c) the measurements are relatively uniformly grey, meaning that any reconstructed sample is unlikely to bear a resemblance to the measured physical sample. Of greatest interest in this chapter are the measurements of (b), in which large-scale pores are resolved, but fine-scale pores are not, leading to a balancing of assertions and contributions between the prior and measurements.



(a) Both scales resolved, $d = 5$ (b) Single scale resolved, $d = 15$ (c) No scales resolved, $d = 50$

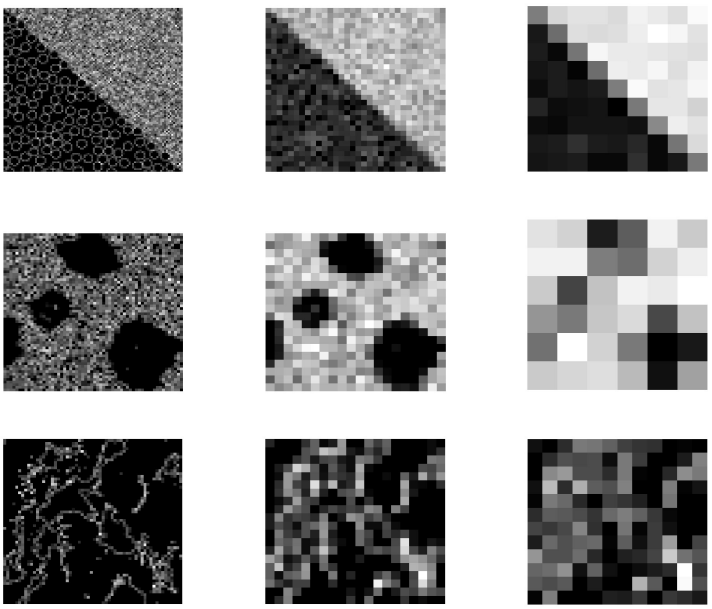
Figure 4.1: Porosity measurements at different resolutions (with downsampling factor d). Particularly in (c) nearly all information of the image structure has been lost, and the distinction of the two phases in the image is not inferrable from the measurements.

Other measurements can be defined which can provide complementary information on fine-scale structures. Consider, for example, a spatially-resolved measurement of the decay of transverse nuclear magnetization, $M_z(t)$ [47]. This measurement is generally described in terms of a multi-exponential distribution of apparent relaxation times T_{2i}

$$\frac{M_z(t)}{M_0} = \sum f_i \exp^{(-t/T_{2i})\ 1} \tag{4.1}$$

where f_i is the volume fraction of fluid relaxing at a rate $1/T_{2i}$. A number of processes can potentially influence magnetization decay in porous media, including bulk and surface-enhanced relaxation, and relaxation due to diffusion in the internal and external gradients. The effects of relaxation due to diffusion can be minimized by a proper choice of the pulse sequence and the effect of bulk relaxation, so that the distribution of surface-enhanced relaxation times T_{2s} may be extracted from the

[1]This equation and some of the text in this section are borrowed from [52] which includes contributions from Professor Marios A. Ioannidis and Professor Paul Fieguth.



(a) $d = 5$ (b) $d = 15$ (c) $d = 50$

Figure 4.2: Simulated surface-to-volume ratio measurements at three resolutions. The measurement is a function of pore structure, even at unresolved scales. In sharp contrast to Figure 4.1, the delineation of the two phases in the top row is clear at even very low resolutions.

distribution of apparent relaxation times. A T_{2s}-distribution may be interpreted as a volume-weighted pore size distribution. That is, each relaxation rate $1/T_{2s}$ corresponds to a different pore length scale ℓ, where ℓ is the ratio of pore volume to pore surface area, V_p/S, of a region of pore space where uniform magnetization is maintained by molecular diffusion over measurement times [4, 48, 53]. MRI is thus able to provide a measurement M_d, reflecting information on the average surface-to-volume ratio within a pixel in 2D or a voxel in 3D. Each datum in measurement M_d reflects the average pore size of a given block of pixels of the underlying high resolution field Z. Such a measurement provides valuable clues with respect to unresolved structures and geometry, as illustrated in Fig. 4.2.

Moreover, as can be seen from the scatter plot in Fig. 4.3 there is no specific linear correlation between the information that the local porosity measurement and the surface-to-volume measurement are providing. In other words, each type of measurement is able to provide independent information on pores and solid structures from another.

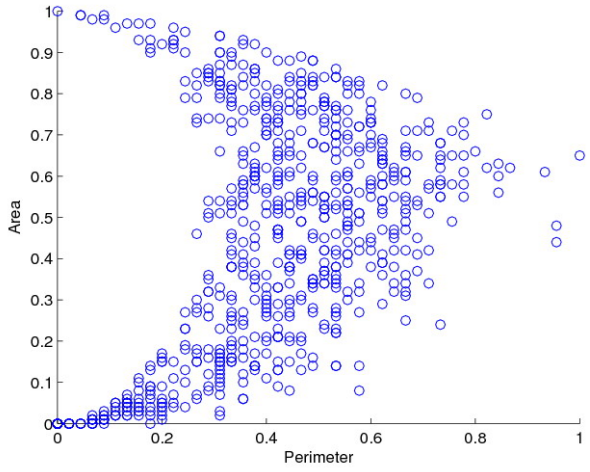


Figure 4.3: Scatter plot of perimeter and area. The plot provides a notion about the correlation between the *area* of pores in 2D (equivalent to the volume of pores in 3D), corresponding to the local porosity measurements, and the *perimeter* of pores in 2D (equivalent to the surface of pores in 3D).

Including both measurements leads to a modification of the constraint term in the posterior energy of Eq. (3.2) as

$$J(Z;M) = J(Z;M_d,M_p) = \|f_d(Z) - M_d\| + \gamma\|f_p(Z) - M_p\| \tag{4.2}$$

where $f_d(\cdot)$ and $f_p(\cdot)$ are the forward models for surface-to-volume ratio and porosity measurements, respectively, and γ is a weighting parameter between the two constraints.

The forward model for obtaining the local porosity measurement, f_p, is based on downsampling the simulated Z according to Eq. 3.5.

The forward model f_d corresponds to the relaxation of T_2. A datum f_d^i in $f_d(Z)$ is defined as

$$f_d^i = \frac{f_p^i}{P(B_i)} \tag{4.3}$$

the ratio of area to perimeter, where the perimeter $P(\cdot)$ is defined based on the fraction of pixels which are at the perimeter of pores in a block of pixels B_i in Z corresponding to f_d^i in the measurement space.

4.3 Posterior Sampling

If we view the measurements as noiseless and the prior energy as a set of empirical hard constraints, then the set of minimal energy configurations becomes

$$\{Z \mid H(Z) = 0 \text{ and } J(Z; M) = 0\} \tag{4.4}$$

That is, the posterior sample is a random Z chosen from the combined null-spaces of the prior model H and measurement model J.

In practice, there are multiple reasons why this hard-constraint perspective becomes impractical:

1. It is possible for the null spaces of H and J to be non-intersecting, meaning that *no* solution exists for Z. In other words, there might be inconsistencies between the information provided by the prior model and the measurements, such that all valid structures in one will be forbidden in the other.

2. An energy function which contains terms which either permit or forbid hypothesized states leads to an energy map with abrupt local minima, which present difficulties to simulated annealing. Continuous energy functions are more likely to lead to robust convergence.

3. Given a finitely-sized sample, it is not possible to infer *rigid* rules regarding the prior model. At best we can infer that certain behaviors are more or less likely to occur, again leading to a continuous definition for the prior energy H.

If the measurements are treated as exact, then our posterior sampling problem is an example of annealing subject to hard constraints, a problem which has previously been investigated [29], in which we seek a random sample from $p(Z)$, subject to precisely matching the measurements:

$$\{Z \ \sim \ p(Z) \mid J(Z; M) = 0 \ \}. \tag{4.5}$$

Since the hard constraint is not easily asserted, in practice this is most easily accomplished within the context of annealing by asserting the hard constraint to an ever increasing degree, annealing over

$$H(Z|M) = H(Z) + \alpha J(Z; M) \tag{4.6}$$

such that $\alpha \to \infty$ as $T \to 0$. The result of this annealing is to produce a posterior sample, forced to be consistent with the measurements, and randomly sampled from the prior within the constrained space.

In practice we have used an exponential cooling schedule for T, starting with a high temperature T and small α, and continuing by decreasing T very slowly with constant or slowly increasing α. The sampling process starts with a purely random binary image, without any special initialization or porosity assertion, since the measurement in the energy function will itself constrain the local porosity during the annealing process. When the measurement constraints are satisfied, the temperature is fixed at $T_c \neq 0$ and the Gibbs sampler continues sampling from the constrained space in Eq. (4.5), at non-zero temperature.

In the more usual event of imprecise measurements, the setting of the relative weight α as a function of annealing iteration is less clear and remains an open problem.

4.4 Results and Evaluation

We have applied constrained sampling on various data shown in Figure 4.4 (a): the small-large circle toy problem, a sample of real vuggy carbonate rock, and two physical models of vuggy carbonate rock made of sintered glass spheres [4]. According to repeated tests, all examples are based on a chord length prior, with the exception of the first sintered spheres, for which a histogram distribution acts as the prior model. In all cases an exponential annealing schedule was used, as follows

$$T_j = T_0 \; b^{(j-1)} \tag{4.7}$$

where j indicates the iteration and T_0 is the starting temperature. The parameter γ in Eq. (4.2) was set to make equal the initial energies of the porosity and surface-to-volume ratio measurements, and the initial value of α in Eq. (4.6) was set to make an initial balance between the prior and measurements terms. The sampling algorithm is converged when then energy does not changed for around 100 consequent iterations. Using MATLAB programming on a machine with 3.0 GHz CPU and 1GB RAM, the reconstruction of 512×512 samples are generated within twenty hours.

A first set of samples, generated by the proposed constrained sampling approach, are shown in Fig. 4.4. In row (d) of this figure we can see a stunning enhancement of the low-resolution measurements of rows (b) and (c). Indeed, (d) reproduces

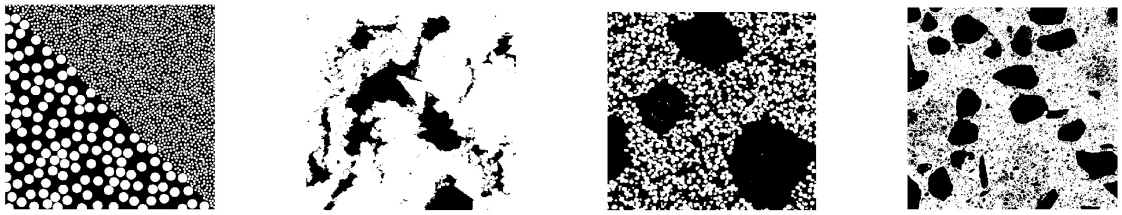
many of the finer-scale features seen in the original images of (a), but which are completely lost in the measurements (b). Although the prior model does not have any specific preference for the two different structures represented in the small-large circle data, the reconstructed sample *does* contain structures at two distinguishable scales due to using the information provided by the measurements. The samples are not perfect, of course, and are particularly limited by the sophistication and quality of the prior model. In particular, our use of a single, stationary prior model has some difficulty in reproducing fine details in the nonstationary fields of the leftmost and rightmost columns in Fig. 4.4.

Fig. 4.5 shows three posterior samples of sintered glass spheres, all three drawn from the same prior model and the same measurements. That the three samples are different, despite being drawn from identical prior and measurements, serves to emphasize the *random* posterior sampling nature of our approach. Features which are resolved by the low resolution measurements, such as large pores, well appeared in their original location where their edged have some perturbations. The unresolved features are constrained by the pattern of low resolution measurements, however the unconstrained degree of freedom are randomly sampled.

To quantify our results we begin with Fig. 4.6, which shows the correlation between the original and reconstructed data. Because a pixel in the middle of a large pore or solid is likely to be the same in the original and reconstructed images, as opposed to a pixel on a pore or solid boundary, the correlation is computed as a function of scale, where the scale of each pixel is defined as the number of scales over which a pixel's value is unchanged under repeated decimation. The resulting figure reflects the degree to which the results in Fig. 4.4 (c) are consistent with the original data. As seen in Fig. 4.6, the correlation is significant below the measured scale, shown as a vertical dotted-line, emphasizing that the structures in the reconstructed results *below* the measurement scale remain consistent with the original data. The MATLAB code in Appendix A shows how the correlation is computed.

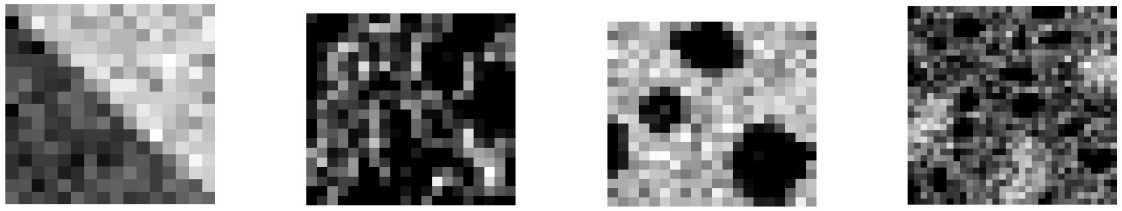
The proposed method has been compared with another method described by Okabe and Blunt [5], which proposes a modified form of posterior sampling, in which areas where measurements are purely white or black (i.e., pore or solid) are preserved, with the remaining values filled by *prior* sampling. Because the fraction of black or white pixels in the measurements is clearly a decreasing function of the decimation factor d, as shown in Fig. 4.7, the performance of this method is likely to suffer as d increases. Indeed, the results in Fig. 4.8 confirm this conjecture, and furthermore show the strength of the proposed posterior sampling approach. Fig. 4.8(c) suffers from blocky structures and pores completely disappearing as they

Small-large circle toy problem | Carbonate rock | Sintered glass spheres (case 1) | Sintered glass spheres (case 2)

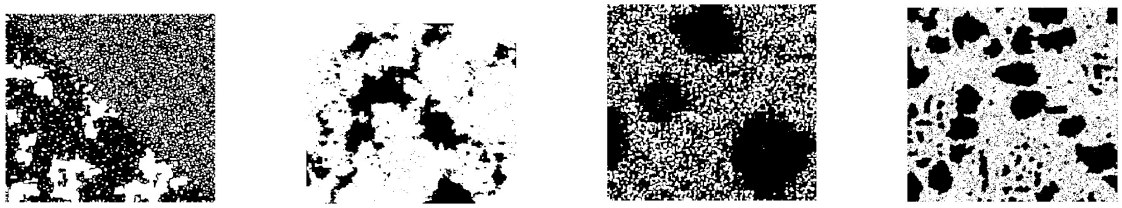


(a) Original media

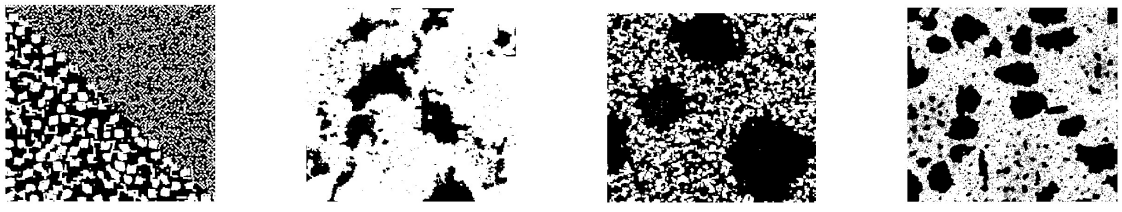
(b) Low resolution porosity measurements, downsampled by $d = 15$ from (a).



(c) Low resolution surface-to-volume ratio measurements



(d) Reconstruction, using measurements in (b)



(f) Reconstruction, using both measurements in (b) and (c)

(See the proceeding page)

Figure 4.4: Artifically reconstructed samples (bottom), using posterior sampling, based on the low-resolution measurements in (b) and (c). The improvement in detail in (f) relative to (b) and (c) is striking. Some of the reconstructed samples (specially the small-large circle example) in (d) do not have the same pore size structures as the original, since the local porosity measurement solely can not provide detailed information on the pore size structures at different scales.

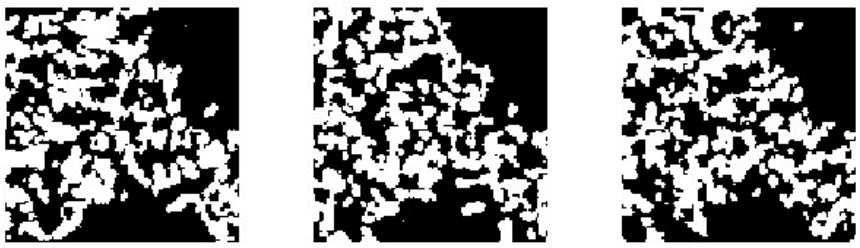


Figure 4.5: Three posterior samples: observe that the large, resolved structures (such as the pore, top-right and bottom) remain unchanged, whereas unresolved details (fine-scale structure) are randomly synthesized. Thus we are constructing multiple samples, all representative of a given measured medium.

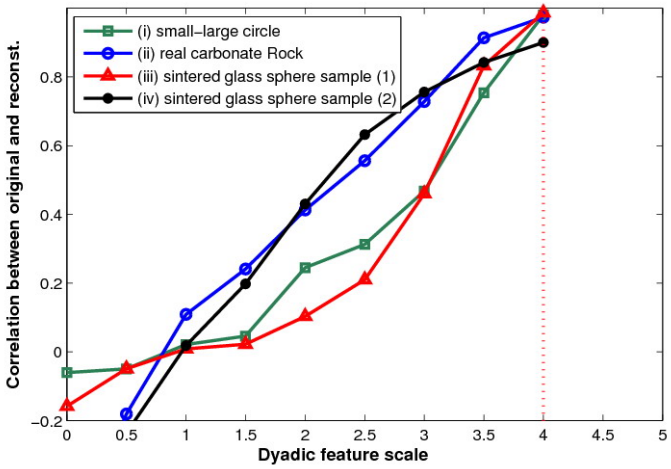


Figure 4.6: The correlation between the reconstruction results in Fig. 4.4(d) and their corresponding original images. The dotted line shows the measured scale. It is clear that the reconstruction is correlated with the true field one or two scales finer than measured. That is, the proposed posterior sampling approach *does* add value below the measured scale; the fine-scale details are not just *random*, rather they are consistent with the measurements, where constrained by the measurements, and consistent only with the prior model where not constrained by the measurements.

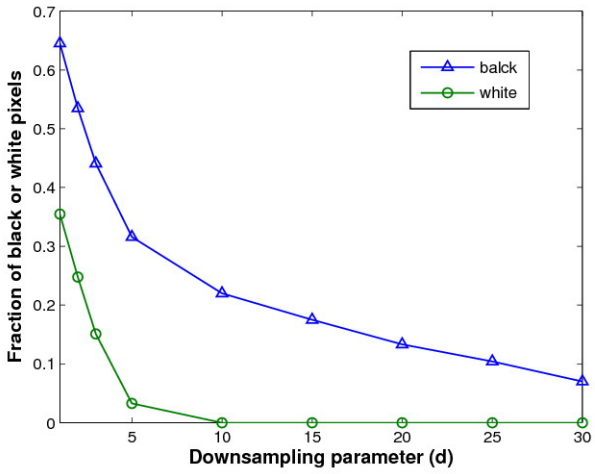


Figure 4.7: Fraction of purely black or purely white pixels in the measurements as a function of downsampling for the sintered glass spheres.

fail to be explicitly resolved in the low-resolution measurements. The key to our results is that a given feature is not either perfectly resolved or fully unresolved, rather the degree to which the measurements resolve a feature lies on a continuum. Even grey, low-resolution measurements offer *some* constraint on the reconstruction, and it is the *interaction* between measurements and prior model which provides the consistent results, specifically with respect to the middle pore in Fig. 4.8(d).

Next, Fig. 4.9 shows the correlation, computed as in Fig. 4.6, between the reconstructed samples and original data for three different approaches:

1. Random sampling (no prior) constrained by the measurements,
2. Prior sampling constrained on pure black and white pixels, as in [5],
3. Our proposed posterior sampling.

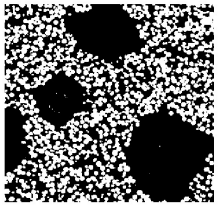
In the first method no prior is involved and the results are generated randomly, only consistent with the measurements. The second method [5] leaves the pure solid and pore areas unchanged in the reconstructed image, while the rest are generated through a prior sampling approach. As is expected, using only measurements (first method) or using uncoupled measurements and prior (second method) lead to reconstructions in which little or no relevant detail is created at scales finer than the measured one, whereas our coupled, measurement-prior posterior sampling approach leads to more strongly correlated details.

A similar conclusion can be reached by examining the overall reconstruction accuracy, measured in terms of Mean Squared Error (MSE), between the original and reconstructed samples. The MSE between true field G and reconstructed field Z, can be computed as

$$MSE = \frac{1}{n^2} \sum_i |g_i - z_i|^2 \tag{4.8}$$

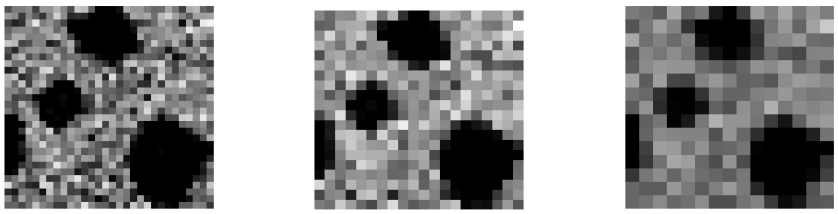
where n is the size of the image. Fig. 4.10 plots the MSE for the three different methods described above as a function of decimation factor d. Compared with the two other approaches, posterior sampling produces results more consistent with the original sample.

To study how much the structures in the results are similar to real porous media, we have evaluated the reconstructed samples under statistical models learnt from real samples, such as histogram and chord length models shown in Eq. (3.4) and (3.3). These models reflect valid statistical structures and features of porous media. The values shown in Table 4.1 in fact, indicate the dissimilarity between the

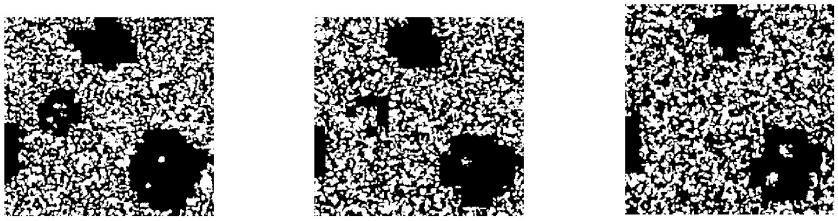


(a) The original data

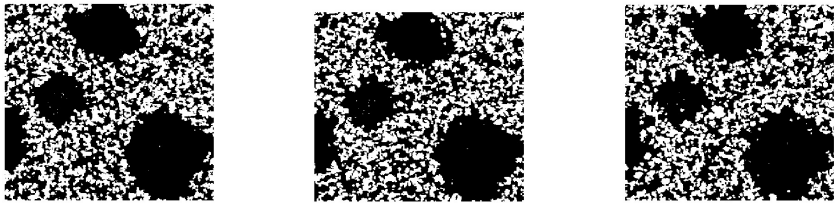
Decimation $d = 10$ Decimation $d = 15$ Decimation $d = 20$



(b) Local porosity measurement at different resolutions

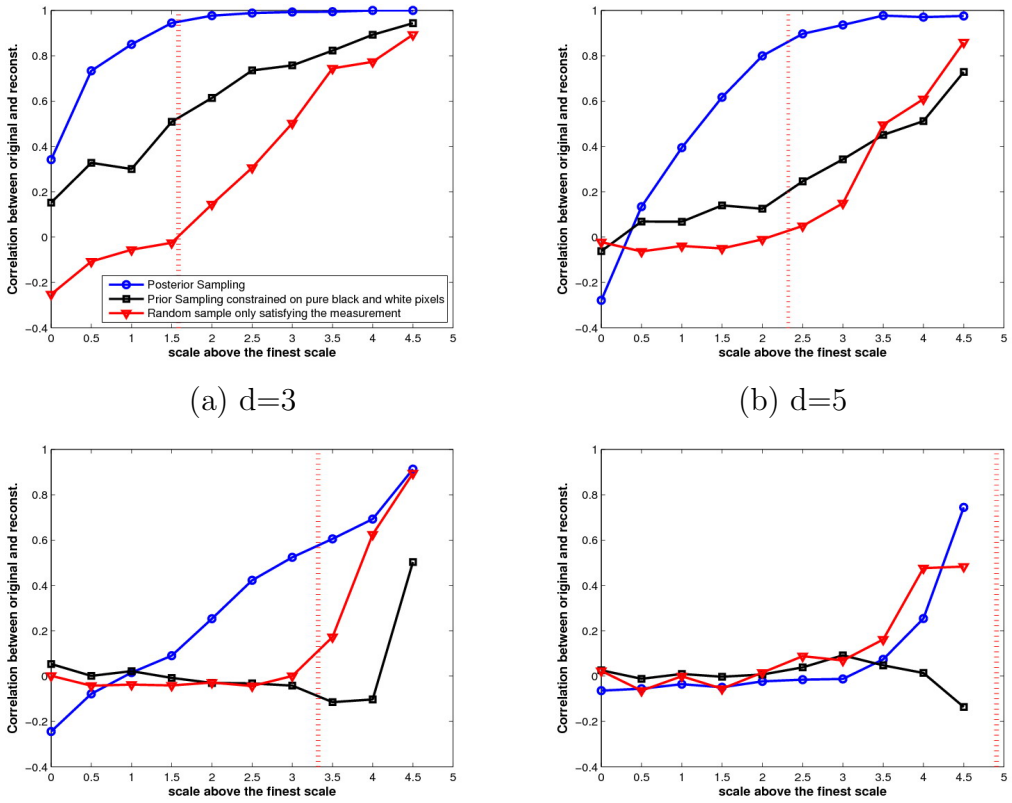


(c) Reconstruction, using prior sampling as in [5]



(d) Reconstruction, using the proposed posterior sampling

Figure 4.8: Reconstructed artificial samples. Because the method of [5] relies on the presence of purely white and black measurements, as d increases the reconstructions suffer from blocky artifacts and lost pores (as in (c)). Posterior sampling, on the other hand, is able to use shades of gray in the measurements as constraints on the random field, and therefore is able to tolerate reduced resolution more gracefully.



(a) d=3 (b) d=5

(c) d=10 (d) d=30

Figure 4.9: Correlation between reconstructed and original data (for sintered glass spheres) comparing our proposed approach (posterior sampling) with two other methods, as a function of scale, for four different values of the decimation parameter d. The dotted line shows the measured scale in each case. The proposed method is more successful in reconstructing structures finer than the measured scale, with the distance between the circled and squared curves essentially representing the improvement of our proposed method over that of [5].

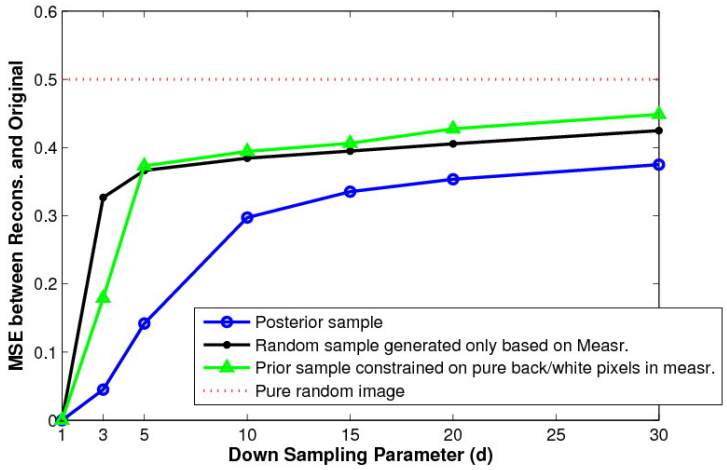


Figure 4.10: Mean-Square Error (MSE) between the reconstructed and true fields as a function of decimation factor d for sintered glass spheres. The four lines show MSE corresponding to a purely random field, a prior-free reconstruction based on measurements only, the prior sample constrained on pure black and white pixels in the measurement [5], and our proposed sample.

Table 4.1: Dissimilarity between the original and reconstructed samples for sintered glass spheres, case 1. Comparing the numbers within each row, we observe that using both measurement lead to smaller numbers, meaning more consistency with the original data.

Measurement / **Prior model**		M_d & M_p	M_d	M_p
chord length model	Horizontal chord	**0.0089**	0.0581	0.0118
	Vertical chord	**0.0112**	0.0724	0.0142
Histogram model	White histogram	**0.000007**	0.00150	0.00013
	Black histogram	**0.000016**	0.00031	0.00003

reconstructed and real samples of porous media in terms of chord length and histogram models. These values are obtained by supplying the reconstructed results into Eq. (3.4) and (3.3) when using different types of measurements in the reconstruction process. As can be seen from this table, when both measurements are involved in the reconstruction process, the dissimilarity decreases, i.e. the results contain more consistent porous media structures and features.

Finally, in Fig. 4.11 we illustrate the application of our proposed approach to

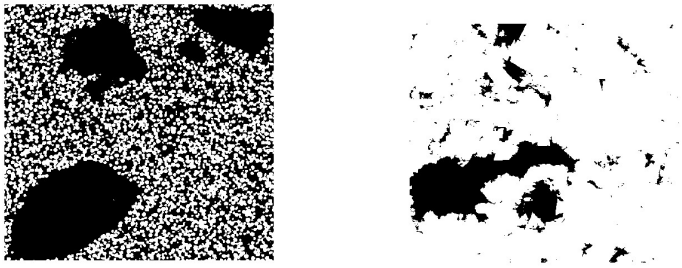
real MRI measurements. The prior model is learned from accurate, high-resolution images, shown in Fig. 4.11(a). However, in contrast with previous experiments, in which we synthesized the measurements from high resolution data, in this case the measurements are actual MRI data [4], taken from a different, but statistically consistent, sample. The resulting reconstructed results in Fig. 4.11(c) are much higher in resolution than the measurements in (b). We can see that the large scale structures in both examples are preserved, and that unresolved small scale structures are mostly consistent with the underlying statistics of (a). Since we do not have the measurements corresponding to S/V for the real samples at this time, the reconstruction is only based on one type of measurement — the porosity measurements shown in (b). We would have obtained a better enhancement in resolution and more consistent structures, if the S/V measurement had been also considered in the reconstruction process. Moreover, due to the significant noise in the measurements we also expect, and see, noisy structures in the reconstructed samples primarily, we hypothesize, because the chosen prior models are insufficiently discriminating.

The computational time and memory usage of the reconstruction process is no more complex than that of other energy-based approaches in the literature. If C_P and C_M indicate the complexity of the prior model and the constraint term J in the energy function, then the computational complexity of our proposed method can be written as $C_P + C_M$, relative to C_M for existing methods. Because common prior models are much more complex than the evaluation of the forward model $f(\cdot)$ in the measurement constraint, in general $C_M \ll C_P$ and therefore $C_P + C_M \approx C_P$.

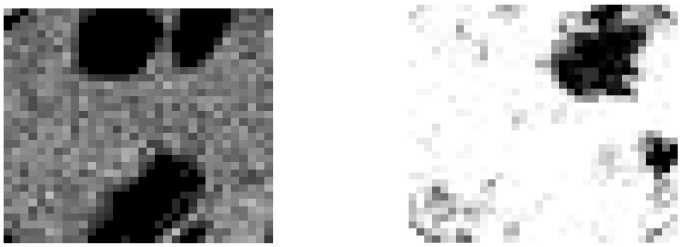
The method proposed by Okabe and Blunt [5] does not anneal over resolved solid and void regions, only reconstructing the part of the image corresponding to grey values in the measurements. If q denotes the fraction of black or white measurements then the computational complexity of the Okabe and Blunt method (C_{OB}) can be written as function of the computational complexity of the proposed method (C_{PS}) as

$$C_{OB} = (1 - q)\ \frac{C_P}{C_P + C_M}\ C_{PS} \tag{4.9}$$

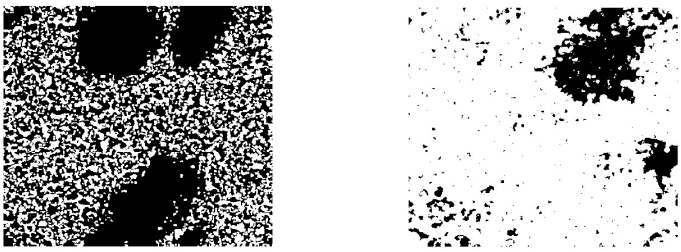
Since q is small compared to the total number of pixels, the term $(1 - q)$ is close to one, leading to $C_{OB} \approx C_{PS}$. Therefore, the computational complexity of the proposed method is comparable to current energy-based methods in the literature in which simulated annealing is used. However, the hierarchical sampling approaches [6, 18, 54] proposed recently are able to decrease the computational complexity caused by annealing process significantly, both in posterior or in prior sampling.



(a) 2D high resolution samples (512×512)



(b) Real 2D local porosity measurements (32×32) [4]



(c) Reconstruction, using porosity measurement (515×512)

Figure 4.11: Reconstruction of two types of porous media using *real* 2D porosity measurements. The prior is learned from a 2D high resolution sample (a), and a single frame of 3D MRI low-resolution sample, shown in (b), is used as the measurement. The posterior samples in (c) are the reconstructed samples of (b), 16 times higher in resolution. The samples in (a) and (b) are not identical, but are assumed to obey the same statistics.

4.5 Summary

In this chapter we proposed a statistical fusion approach based on posterior sampling for two-scale porous media reconstruction, in particular, with problems having one resolved and one unresolved scale [52, 55, 56]. The key to our approach is the simultaneous assertion of prior and measurement constraints, which leads to superior reconstructions, possessing useful details at scales finer than that of the measurements. An extension of the proposed approach to three dimensions should be possible in principle. The method is likely to be limited by computational complexity, unless based on a hierarchical approach which is going to be discussed in Chapter 6.

Chapter 5

Annealing with Constraints

Generating artificial sample, solely based on a prior model is widely used in the literature for the purpose of porous media reconstruction [3, 6, 17, 18]. In this method the prior energy is minimized using simulated annealing as discussed in Section 2.4.3 , equivalent to maximizing the prior probability distribution. However, as will be shown in this chapter, the samples generated using this method can not represent the variability of the pore/solid structures in a given porous material, since a single prior model can only assert a limited set of plausible or implausible pore/solid structures. In this chapter, a constrained sampling approach is proposed to overcome the ill-representation of the prior model. Although simulated annealing is used to generate artificial samples, the proposed approach is based on sampling from the constrained Gibbs probability distribution rather than maximizing the probability distribution.

5.1 Problem Statement

Porous media reconstruction using Gibbs sampler along with simulated annealing (a widely used method in the literature [3, 9, 10] leads to minimizing the energy function. Simulated annealing, as explained in Section 2.4.3, is an optimization method which can also be done under deterministic constraints depending on the problem at hand. Maximizing the probability distribution of an image using simulated annealing leads to the most probable realization which can not represent the variability of different structures in the data at different scales. Moreover, as discussed in Section 3.2.1 and 3.3 it is not possible to find a true prior model to represent all aspects and variations of structures in a given porous material, therefore

prior sampling solely based on a single prior model may lead to artificial samples with structures at a single scale .To cover the ill-representation of the prior model, we propose a constraint term in the model and also propose a constrained sampling approach based on simulated annealing. In summary

> *We propose to consider a regularization constraint in the energy function, and use a constrained sampling approach, rather than an optimization approach, to generate artificial samples of porous media. The challenge here is how to define and incorporate the constraint in the energy function and how to choose the temperature T and balance between the prior and the constraint to generate the sample.*

In the following section we study the variability of different structures of porous media images at different scales.

5.2 Variability of the Samples

The artificial samples generated from the prior Gibbs probability distribution, Eq. (2.12), are required to represent the variability in the training data S. The variability of the training data can be studied in terms of different statistical descriptors, as the ones discussed in Section 3.2. We consider chord-length distribution function, described in Section 3.2, which is a reasonably widely used probability function in porous media reconstruction [3], to study the variability of the training data, although other types of models can be used for this purpose.

As can be seen from Fig. 5.1, a high resolution sample of porous media, used as the training sample in the reconstruction process, has various structures at different scales.

To study the variations in S at different scales, multiple random s_k – truncated images from S at size $k \times k$ — have been considered. Each s_k has its corresponding chord-length distribution leading to different energy function. In Fig. 5.2 panel (a), superimposition of various chord-length distribution for different s_k is shown, and panel (b) shows the variability of chord-length distribution for different k. This variability is due to having different structures at various scales in the training sample S. As can be seen from this figure, smaller k leads to more variability.

The annealing process used in the literature to generate artificial samples, starts at high temperature, where the probability distribution $p(\cdot)$ is only a weak function

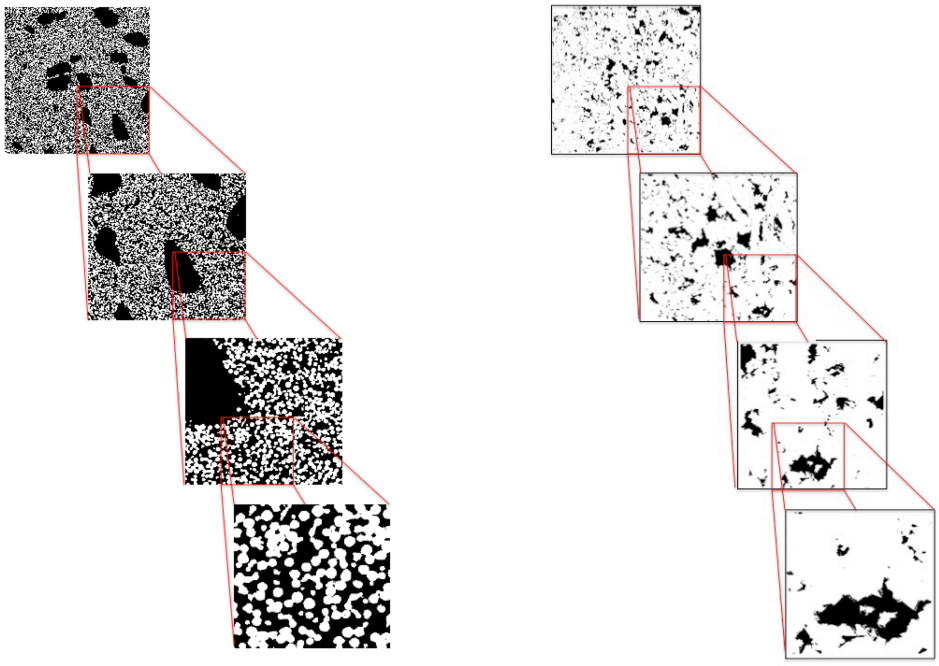
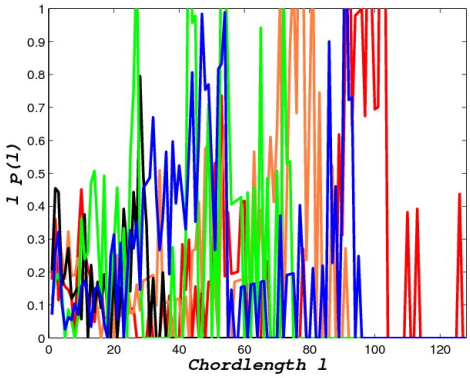
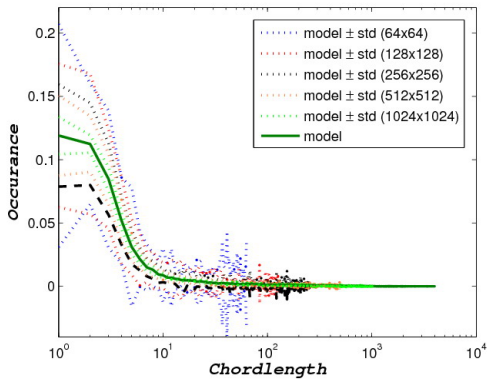


Figure 5.1: Two examples of porous media with different structures at different scales.



(a) Superimposed chordlength distributions for multiple samples.



(b) Variations in the training sample S at different scales (k).

Figure 5.2: Variation of the model with respect to the size of the training sample s_k. (a) shows a set of superimposed chordlength distributions taken from different parts of a large image shown in Fig. 5.1. Note the considerable variation between samples. For (b) we have considered a set of images of size $k \times k$ ($k = 64, 128, 256, 512, 1024$), and inferred the average chordlength probability and its variability as a function of k. Clearly as k decreases the variability increases.

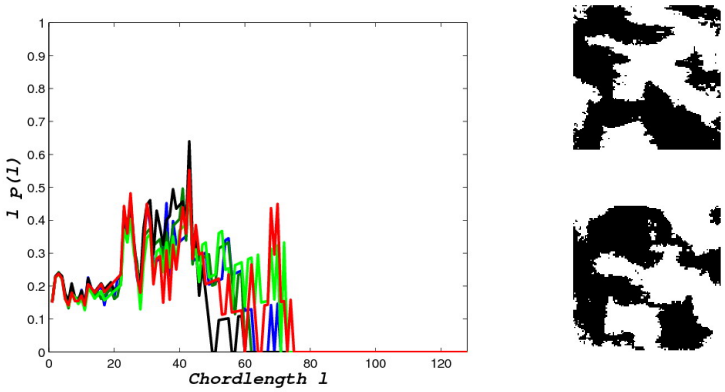


Figure 5.3: Superimposed chordlength distributions resulting from annealing down to $T = 0$, with two example images shown in the right. The variation between artificial images is smaller than the original variation shown in the Fig. 5.2 (a).

of Z, thus Z is relatively unconstrained, and as T decreases the system is driven to lower energy until the minimum energy, the most probable Z maximizing $p(Z)$, is obtained. However, the images generated by this method can not represent the original variability in the training sample shown in Fig. 5.2, since T is decreasing down to zero and consequently the probability distribution $p(\cdot)$ is maximized. As shown in Fig. 5.3, the superimposed chord-length distribution of the generated images through this method are almost the same and do not contain the required variability in comparison to the original variability shown in Fig. 5.2 (a).

Theoretically, for Gibbs sampler to generate uniformly distributed random samples from Gibbs probability distribution along with simulated annealing, the relaxation parameter T should be decreased down to a non-zero finite value such as $T_f \neq 0$, and the rate of decrease should be logarithmic [34]. However, we do not know at what finite value of T the Gibbs sampler generates independent random sample from the chordlength distribution. Even if T_f can be determined, the logarithmic annealing schedule causing high computational time, should be used to guarantee generating independent random samples, and using fast non-logarithmic schedules leads to *quenching* (rapid cooling), instead of annealing. The small variability in Fig. 5.3 is due to quenching rather than a typical variability in typical random samples. Quenching assumes upon decreasing T faster that the logarithmic schedule, since logarithmic schedule involves very large computational time.

Moreover, even if we anneal the energy function down to small enough $T_f \neq 0$ (using logarithmic or non-logarithmic schedule), although large scale structures are

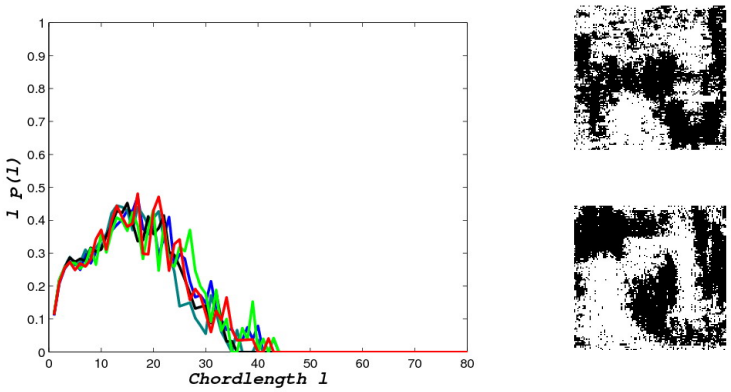


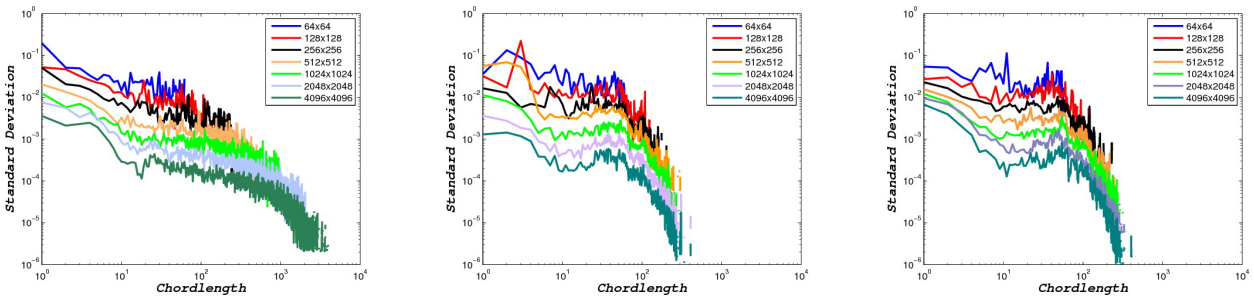
Figure 5.4: Superimposed chordlength distributions of samples generated by annealing down to $T_f \neq 0$. Although the variation in small scale structures (small chordlength) is 21% more than the variation of the results generated by annealing down to $T = 0$ in Fig. 5.3, the samples contain illegal morphologies and structures.

presented and there is more variations in the samples (since $p(\cdot)$ is not maximized), the result contains illegal small scale morphologies and structures. Fig. 5.4 shows superimposed chordlength distribution of the artificial samples generated by annealing down to non-zero temperature T_f. As can be seen from this figure, the variability in small scale structures of the samples generated by annealing down to $T_f \neq 0$ is 21% more than the variability of the results generated by annealing down to $T = 0$.

However, the variability is not inherent in porous media, rather it is a function of scale. As can be seen from Fig. 5.2 (b), as k gets larger, the degree of variability decreases. In general, for $1/f$ (power law, self-similar) [57] processes, there is almost as much as variability in fine scale as in coarse scale, while for ergodic processes (as in some of the porous materials), there is one scale from which the statistics is stationary. Studying the variability in the training samples at different scales, we can see from Fig. 5.5 that standard deviation (std) of a set of chordlength distributions obtained from multiple s_k decreases as k increases. Thus, the random field is ergodic from one specific scale (k^*). Therefore to capture the variability in the training sample a possible approach is to synthesize the image at scale k^* – the scale that the ergodicity of the random field is observable.

Thus, for generating random samples, one can

- synthesize an image at size $k^* \times k^*$ by simulated annealing. Since k^* is very large, synthesizing an image at that size requires a very large computational



(a) Carbonate rock (b) Sintered glass spheres with vugs (large pores) (c) Sintered glass spheres without vugs

Figure 5.5: Variation of the chordlength model with respect to the size of the training sample for different types of porous media images.

time.

- consider a smaller scale $k < k^*$, and synthesize the image at that scale by simulated annealing. As discussed earlier and shown in Fig. 5.3, the results do not represent required variations in porous media.
- down-sample the training sample and synthesize at smaller scale by simulated annealing. By down-sampling the image, we will loose small-scale morphologies information, therefore we can not reconstruct the small scale structures.
- use variable models at different scales and synthesize each by simulated annealing. This makes the problem more complicated, and needs an annealing method for variable models.
- use fixed model and synthesize an image at size $k < k^*$ by *constrained annealing*.

Through the constrained sampling approach we propose to change the energy function by adding a constraint and using constrained annealing to generate samples form the model. More details on this approach follow in the next section.

5.3 Sampling with Constraints

To generate typical random samples from Gibbs probability distribution, we propose to

- change the energy function to a constrained energy function,
- generate samples from the constrained energy function using constrained annealing.

The constraint term in the energy function can be measurement, as discussed in Chapter 4, or another prior model which avoids generating invalid small scale structures observed in the non-zero temperature annealing. The constrained energy function is defined as

$$H_{cons}(Z) = H(Z) + \alpha J(Z) \tag{5.1}$$

where $J(Z)$ is the constraint and α is a parameter controlling the contribution of constraint in the energy function. The constraint term is considered to be another model learnt from the training sample. More specifically,

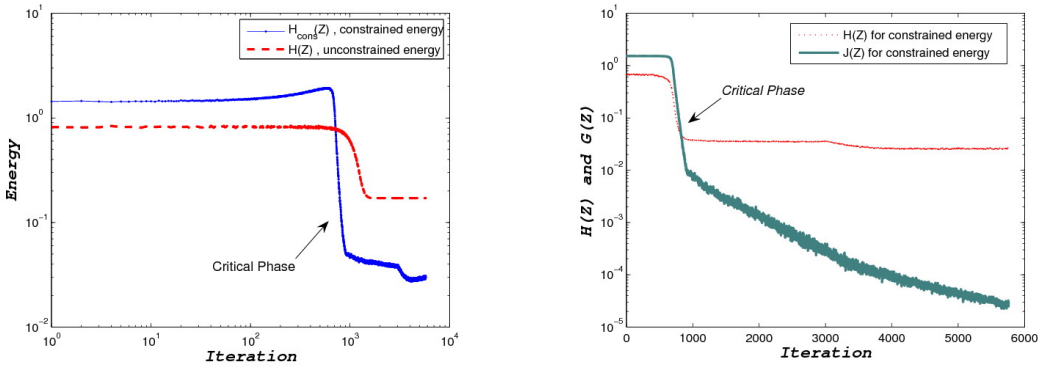
$$H_{cons}(Z) = H_c^1(Z) + H_c^2(Z) + \alpha H_h(Z) \tag{5.2}$$

where H_c^1 and H_c^2 are the chord-length distributions for the pore and solid phases, respectively, as defined in Eq. (3.4), and H_h is the histogram distribution as defined in Eq. (3.3). The energy function defined in Eq. (5.1) does not contain any term related to the measurement. However, in the reconstruction from a low resolution measured sample, one can add another term the same as discussed in Chapter 4, as the measurement constraint.

Having the constrained energy function defined in Eq. (5.2), we need to generate samples from the model. Basically, as discussed in Section 3.3 and in [29] one should fix $T \neq 0$ and increase α up to infinity to satisfy the constraint. However, we do not know what specific value for T leads to porous media samples.

Therefore, we propose to start with large T and small α, and decrease T while increasing α very slowly. This annealing process has a critical phase in which the large-scale structures are generated. Up to this phase both terms in the energy function contribute to assert valid structures in the sample. As shown in Fig. 5.6, by critical phase we mean when a steep decrease happens in the energy function. The energy level for the following methods has been studied in Fig. 5.6

1. annealing the unconstrained energy function H down to $T = 0$,
2. annealing the constrained energy function H_{cons} down to T_f while increasing α up to infinity.



(a) Energy level for constrained and unconstrained energy

(b) Prior and constrained term in the constrained energy function

Figure 5.6: How the energy function changes as a function of iteration for two approaches: unconstrained and constrained annealing. In both annealing approaches there is a steep decrease in the energy function, that we name as the critical phase. In panel (a) the energy level for both approaches is illustrated. Panel (b) shows how both terms in the constrained energy (H_{cons}) changes for the proposed approach.

Fig. 5.7 shows the samples generated from constrained energy H_{cons} and their variations, right after the critical phase. As can be seen from the superimposed chord-length distribution in Fig. 5.7, the variation between samples is more than the variations between images shown in Fig. 5.3 generated from unconstrained annealing.

After the critical phase, we set T to be fixed (T_f), and by continuing increasing α let the constraint term change the small scale structures. According to Fig. 5.3 sampling at T_f from the *unconstrained* energy function in Eq. (2.12) does not lead to a result with valid structures, or we may say that it does not generate porous media samples. With the proposed method, we are sampling at T_f, while the constrained term $J(\cdot)$ contributes in the process from the beginning and also after T is fixed. Thus, it can change the small scale structures into the valid structures. Therefore, by the proposed method we end up sampling from the constrained space

$$\{Z|Z \in \Omega, \quad J(Z) = 0\}. \tag{5.3}$$

Fig. 5.8 shows the variability of the samples generated by the proposed method. We can see that the variability is 39% more than the one in Fig. 5.3.

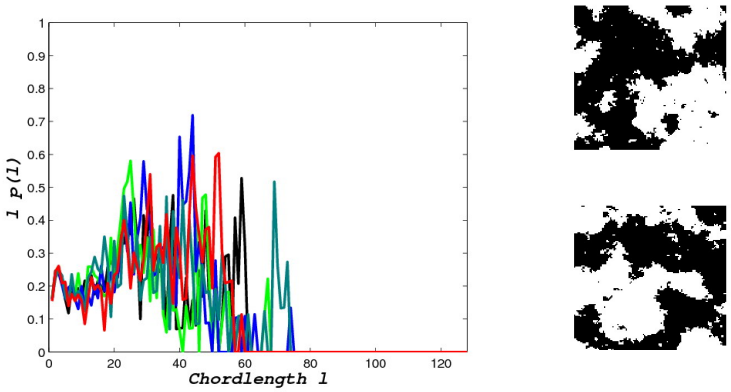


Figure 5.7: Constrained sampling by annealing down to finite T and increasing α up to $\alpha_f \neq \infty$. T_f and α_f are set according to the critical phase(Fig. 5.6), when a steep decrease happens in the energy function. Although the variations in the samples are 22% more than the unconstrained annealing shown in Fig. 5.3, the samples still contain noisy and invalid structures.

The proposed approach can be generalized to have more than one constraint in the energy function and changing the parameter for each constraint to control their degree of contribution in the whole process. Although, we can synthesize and reconstruct porous media at smaller scale while having more variations in the samples, for larger image synthesis we still cope with large computational time, and the original variability in the training sample has not been reached completely.

The proposed method is describe in Algorithm 4. The rate of decrease for T is given in Eq. 2.29 and the increase rate for α is as follows

$$\alpha_k = \alpha_0 a^{k-1} \tag{5.4}$$

where a is close enough to 1, such that $a = 1 + \epsilon$, for small $\epsilon > 0$.

5.4 Results and Evaluation

We have considered chordlength distribution for the prior term $H(Z)$ and histogram model as the constraint term $J(Z)$ in the constrained energy function.

Fig. 5.9 shows 256×256 samples generated using constrained sampling. We have evaluated the artificial samples in terms of the following criteria

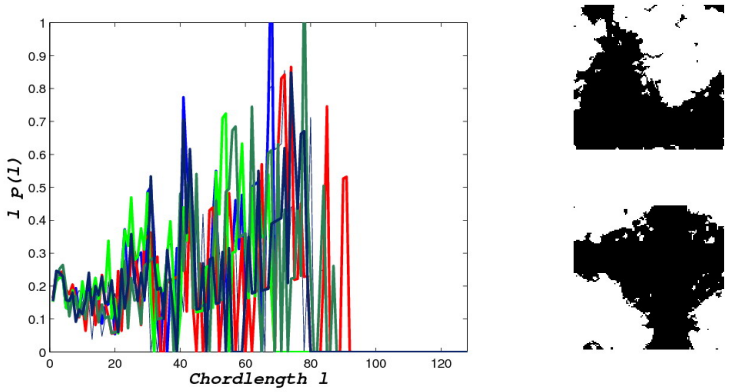


Figure 5.8: Constrained sampling by annealing down to finite T_f and increasing α up to infinity. T_f is chosen according to the critical phase, and the algorithm will stop when the constraint in the energy function is satisfied, i.e. $J(Z) = 0$. The variations in the samples as shown in the left panel are 39% more than the variations in Fig. 5.3 obtained by annealing down to $T = 0$.

- How consistent the structures and morphologies presented in the results are with porous media
- How close the results are to a typical random sample

For the first criterion, we have considered two different porous media models learnt from the training sample s_k at size 128×128 and evaluate the results in terms of those models. We have evaluated the samples obtained from constrained sampling and unconstrained annealing in terms of histogram model. The evaluation shows that the error for unconstrained annealing (annealing the unconstrained energy down to $T = 0$) in terms of histogram model is 17%, while for constrained sampling it is 3%. The error for unconstrained annealing in terms of chordlength model is 0.74%, while for constrained sampling it is 0.06%. Fig. 5.10 shows how much the reconstructed samples are compatible with the chordlength model learned from the original training sample.

For the second criterion, we study the variations in different samples generated by constrained sampling and compare that with the following methods:

- Annealing the unconstrained energy down to $T_f \neq 0$
- Annealing the unconstrained energy down to $T = 0$

(a) Samples generated by annealing the unconstrained energy down to $T = 0$

(b) Samples generated by constrained sampling

Figure 5.9: Samples generated using annealing down to $T = 0$, and constrained sampling.

Algorithm 4 Constrained Sampling

Start with an initial random field $Z^{(0)}$
Initialize T and α
while $J(Z) \neq 0$ **do**
    Update $Z^{(i+1)}$ from $Z^{(i)}$ according to Gibbs sampler
    if $T < T_f$ (the critical phase has not reached) **then**
        Decrease T according to exponential schedule
    else
        Let $T \leftarrow T_f$
    end if
    Increase α according to increase rate
    $i \leftarrow i + 1$
end while

The variation in samples generated by constrained sampling is more than the other methods, as confirmed in Fig. 5.11. Although there is variability in the results generated by annealing the unconstrained energy down to $T = 0$, the variability is due to quenching rather than sampling. The proposed approach generate samples with 39% improvement in the variability.

5.5 Summary

In this chapter, an approach for sampling from Gibbs probability distribution is proposed which is based on generating samples from the constrained energy function [58]. The previous methods in the literature used for porous media reconstruction are based on maximizing the probability distribution by annealing down to zero temperature, since sampling at finite, non-zero temperature leads to results containing illegal structures and morphologies. These methods generate the most probable realizations which are different from a typical random sample of Gibbs probability distribution, and they can not reflect the scale-to-scale variations. In the proposed method a constraint term is added to the energy function to enable sampling at finite, non-zero temperature while at the same time the generated samples do not contain invalid, illegal structures. According to the evaluation results, not only does the constrained sampling approach generate samples with 39% more variation than the other methods which are based on annealing down to zero temperature, but also the samples are almost ten times more consistent with the original real samples, than the ones generated by other methods.

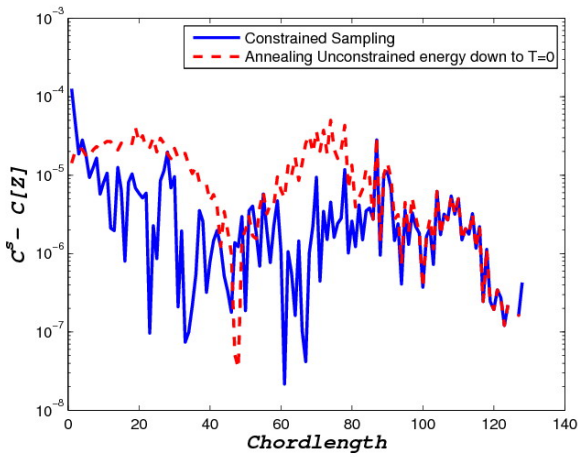


Figure 5.10: How much the results are inconsistent with the original sample in terms of chordlength model. The solid and dashed lines show dissimilarity to the chordlength model for the constrained sampling and unconstrained annealing down to $T = 0$, respectively. As the solid-line is closer to zero, the constrained sampling method generates samples which are more similar to porous media samples in terms of chordlength model, than the other method.

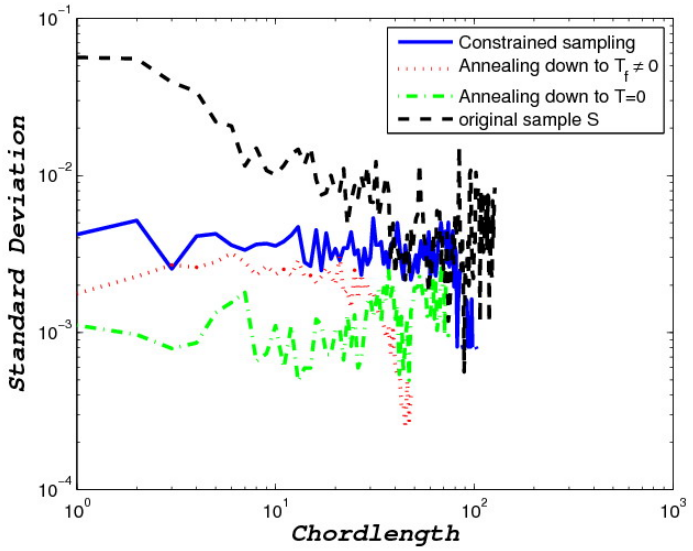


Figure 5.11: Comparing the variations in the results in terms of standard deviation between different methods .The standard deviation of the samples generated by each method is illustrated and compared to the standard deviation of the original training sample. The standard deviation of constrained sampling is closer to the original than the others.

Chapter 6

Hierarchical Posterior Sampling

The large computational time and memory usage of the annealing process mainly caused by decreasing the temperature very slowly, is a challenge in reconstructing large samples of porous media. The hierarchical sampling approaches proposed recently for binary image synthesis [6, 18] are able to overcome this challenge, although these methods are based on prior sampling and no measurement is involved in the process. In this chapter a hierarchical posterior sampling approach is proposed in which the measurement is combined with the prior model at each scale. The main contribution of this chapter is how to define a model to relate the continuous-value low resolution measurement (coarse measurement) to a discrete-value unknown at an intermediate scale coarser than the finest scale in the hierarchical framework.

6.1 Problem Formulation

Problem Statement: As proposed in Chapter 3 the sampling approach for generating porous media samples is based on Gibbs sampler along with simulated annealing. The proposed annealing process is a slow convergence process, and it will end-up with a posterior sample only if the temperature decreases very slowly, based on an appropriate annealing schedule [29, 59]. Reconstruction of *large* 2D samples of scientific images becomes challenging due to the large computational time and memory usage of the annealing process. The challenge turns out to be a barrier in 3D reconstruction of large volumes (usually larger than 256×256) since we are coping with a huge amount of data to store and to reconstruct.

Hierarchical sampling approaches for binary image reconstruction, proposed recently [6, 18], for porous media reconstruction, are able to decrease the computational time significantly in reconstructing large samples. These approaches, developed mainly for binary random fields, decompose the reconstruction problem into several scales, and then choose different strategies to limit the number of data points which are going to be reconstructed at each scale. However, these methods are based on prior sampling, therefore no measurement is involved in the sampling process. Here we propose a hierarchical approach for posterior sampling, i.e. when the measurements are also included in the reconstruction task.

> *Using a ternary state space for the random fields at the coarse scale, the challenge here is how to relate the continuous-value measurement with a discrete-value ternary random field at an intermediate coarse scale in the hierarchical model. We propose a measurement model describing how the measurement is related to an unknown at each scale, and then generate posterior samples based on hierarchical posterior sampling approach.*

Before describing the proposed measurement model, we describe a general hierarchical framework that can be applied for discrete-state random field reconstruction.

6.2 Hierarchical Framework

In this section a general framework for hierarchical approaches that can be applied in porous media reconstruction is discussed. The hierarchical framework is discussed in terms of three main aspects

- hierarchical model: how to define the inter-scale and spatial relationship and how to include the measurement in the model
- state space representation: how the state space of the random fields at each scale is defined
- algorithm: how to generate a posterior sample given a hierarchical framework

Each of these aspects are described in the following three sections and at the end of each section the relationship of each aspect with the framework that is used in this Chapter is discussed.

6.2.1 Hierarchical model

The hierarchical model that is used here is based on the graph-based representation of random fields as described in Section 2.5. More specifically, the hierarchical model considers both spatial and scale Markovianity (the Markov relationship between two consequent scales), therefore it is similar to the model-based hierarchical approaches described in Section 2.5.3. Each scale is represented by a single Markov model similar to the global Markov model.

The spatial Markovianity of the random fields at each scale implies a relationship between a site z_i^k at scale k and its neighbors $\mathcal{N}_k(i)$ at that scale. Therefore, the conditional probability of a pixel z_i^k given its neighbors is

$$p(z_i^k = z^* | \mathcal{N}_k(i)) \tag{6.1}$$

where $z^* \in \Lambda^k$, the state space of scale k. The neighborhood system can encompass the inter-scale relationship as well

$$\boldsymbol{\mathcal{N}}_k(i) = \mathcal{N}_k(i) \cup \Phi_k(i) \cup \phi_k(i) \tag{6.2}$$

where Φ_k and ϕ_k represent the parents and the children of the ith site at scale k, respectively. A more generalized approach is to extend the neighborhood to include the grand-parents and other siblings as well,

$$\boldsymbol{\mathcal{N}}_k(i) = \mathcal{N}_k(i) \ \cup \ \bigcup_{\ell=1}^{u} \Phi_k^{\ell}(i) \ \cup \ \bigcup_{\omega=1}^{f} \phi_k^{\omega}(i) \tag{6.3}$$

where $\Phi_k^{\ell}(i)$ stores the parents of site i at the ℓth coarser scale than the current scale k, and $\phi_k^{\omega}(i)$ indicates the siblings of site i at the ωth finer scale than the current scale k. It is obvious that $1 \leq u < k$ and $k < f \leq n$.

The hierarchical model used in this chapter is the same as the model proposed in [18], and it considers the spatial neighbors, $\mathcal{N}_k(i)$, and the parents at one scale coarser than the kth scale, i.e. $\Phi_k(i)$ as the neighbors of a given site i at scale k. Therefore the neighborhood system that we consider in this chapter is

$$\boldsymbol{\mathcal{N}}_k(i) = \mathcal{N}_k(i) \cup \Phi_k(i). \tag{6.4}$$

In order to include the measurement in the hierarchical model we need a measurement model describing how the measurement interrelate with an unknown Z^k at scale k. In Section 6.3.2 a novel measurement model is proposed for this purpose.

6.2.2 State space representation

The state space of the porous media sample at the finest scale is a binary space: $\{0, 1\}$. The question is how to define the state space for the scales coarser than the finest scale. One possible approach is to consider a binary state space for the coarse scales as well, as proposed in [6]. Considering a binary state space for intermediate coarse scales, there is an ambiguity in what a value 0 or 1 at a given coarse scale means at the finest scale. In other words, in order to arrive at a proper binary pattern at the finest scale, how a 0 or 1 value at a coarse scale inter-relates with 0 and 1 at the finest scale. Ternary state space representation [18] is able to overcome the ambiguity of the relationship between binary values at different scales, to some extent. Rather than considering 0 and 1, the ternary state space representation extends the state space at the intermediate coarse scale to $\{0, 0.5, 1\}$.

Based on the hierarchical framework proposed in [18], the ternary random field representation at each scale and the parent-child relationship described in Eq. (6.4) implies that each scale is rigidly constrained by the parents at the previous scale. More specifically, the sites with black or white parents are frozen as black or white, and only the sites with gray parents can change during the sampling process at a given scale to generate a new sample. Therefore,

$$p(z_i^k = z^* | \mathcal{N}_k(i),\ \Phi_k(i) = z^*) = 1, \qquad when\ z^* \in \{0, 1\} \tag{6.5}$$

and when $z^* = 0.5$ the above probability only depends on the spatial neighbors of z_i^k, i.e.

$$p(z_i^k = z^* | \mathcal{N}_k(i),\ \Phi_k(i) = z^*) = p(z_i^k = z^* | \mathcal{N}_k(i)) \leq 1, \qquad when\ z^* = 0.5. \tag{6.6}$$

In this chapter a ternary state space and the parent-child relationship described in Eq. (6.5) and (6.6) is used. It is notable that the value 0.5 at a coarse scale is only a representation of uncertainty and can not be counted as a real value compared to 0 counted as black, or 1 counted for white. Such uncertainty causes a challenge in relating the measurement with the random field at each scale. The measurement model proposed in Section 6.3.2 is able to overcome this challenge.

6.2.3 Algorithm

Most common hierarchical sampling/estimation algorithms in literature are based on a strategy in which stochastic optimization algorithms such as simulated annealing along with MCMC methods, are used and the results are generated scale-by-scale. Then the sample/estimate generated at each scale is directly or indirectly

propagated into the next finer scales. According to this strategy the coarse scale data can influence the fine scale information during the annealing process, however the fine scale information is not influencing the coarse scale during annealing, since the coarse scale is already frozen. A more general strategy is to anneal over multiple scales simultaneously, to allow with more flexibility at each scale.

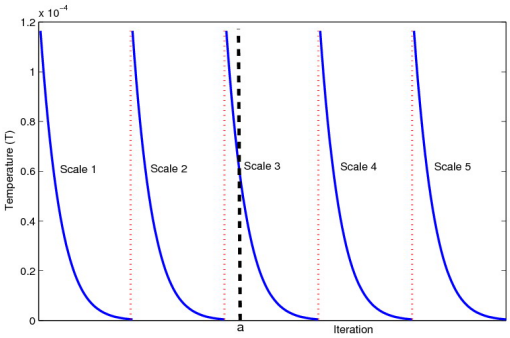
For instance, by using scale-by-scale annealing and a parent-child relationship described in Eq. (6.5), a frozen black site at a coarse scale k leads to a large frozen black area at the scales finer than k, which are constrained by the frozen coarse scale. However, a non-frozen coarse scale causes more flexibility at the fine scale and lets the coarse and fine scales influence back and force during the annealing process. Figure 6.1 shows a schematic representation of the scale-by-scale and simultaneous annealing strategy. According to (ii) in this figure the coarser scales than scale k are not frozen, rather they are *warm*, such that some changes at the scales finer than k can influence the scales coarser than k. An example of the "hot", "warm" and "cold" temperature is shown in Fig. 2.4. In the case of "warm" temperature, the structures are not totally frozen and the small to medium scale structures have a greater degree of freedom to change than the case with the "cold" temperature.

The degree of change in the scales coarser than a given scale is still an open problem and needs to be defined. The main focus of this chapter is how to incorporate the measurement in the hierarchical model, and it is independent of the annealing strategy. In this chapter the scale-by-scale annealing strategy is considered, however, the proposed approach can be also generalized easily to the simultaneous annealing strategy.

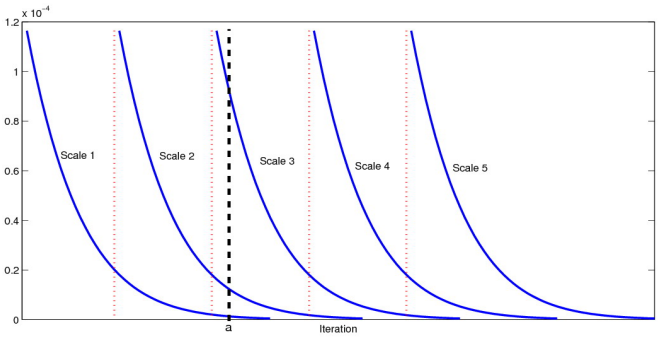
6.3 Hierarchical Posterior Model

Based on the discussion in the previous section, the hierarchical framework that we are using in this chapter considers the spatial and scale Markovianity in the hierarchical model, a ternary state space representation for the coarse scales and a scale-by-scale annealing strategy. The main challenge here is how to define the posterior model at each scale based on the continuous-valued measurements.

The hierarchical posterior model at each scale consists of two parts: the prior model and measurement model. The prior model is characterized by the prior energy which reflects the spatial Markovianity at each scale. In most of the model-based hierarchical models the measurements are estimated at each scale using wavelet or other heuristic multi-resolution decomposition methods. Here, we do not estimate the measurement at each scale, rather we define a measurement model



(i) scale-by-scale annealing strategy



(ii) Simultaneous annealing strategy

Figure 6.1: Temperature profile for the scale-by-scale and the simultaneous annealing strategy. In (ii), the curves for the temperature profile overlap at a given time a, meaning that at a given time a there are multiple scales with non-zero temperature, while in (i) only a single scale is annealed in an arbitrary iteration a.

coupled with the prior model. The following sections describe how the prior and measurement models are defined in the proposed hierarchical framework.

6.3.1 Prior model

According to [18] rather than having binary random field at each scale, a ternary random field is considered. The sample at the finest scale, Z^n, is a binary image which demonstrate the reconstructed image, and at the coarsest scale, Z^0, is in the same scale as the measurement scale. In a ternary random field representation, a site can have three possible value: black, gray, white. According to the graph-based model representation shown in Fig. 2.7 in Section 2.5, given n different scales, a black, white or gray pixel at the kth coarse scale ($k = 0, 1, \ldots, n$) represents a $2^{(n-k)} \times 2^{(n-k)}$ block of black white or gray pixels, respectively, at the finest scale.

Here the histogram model explained in Section 3.2 is considered as the prior model, although other types of prior models can be used as well. Based on the histogram model, for a 3×3 block as the second order spatial neighborhood structure, there are 3^9 possible configurations when we are considering a ternary random field. The prior energy at scale k when using the histogram model is defined as

$$H^k(Z^k) = \sum_{c=1}^{3^9} \frac{|\bar{h}^k(c) - h^k(Z^k; c)|^2}{y^k(c) + \epsilon} \tag{6.7}$$

where $\bar{h}^k$ is the learned histogram distribution and h^k is the observed histogram distribution of a simulated random field Z^k, at scale k. The term y^k is the variance for each histogram entry, to account for sample variation at each scale. A small constant ϵ is introduced to avoid divisions by zero, especially in the comparatively common case of unobserved configurations k corresponding to $h^k(c) = 0$.

The parameters of the prior model are learned from the high resolution training data. The training data at each scale is obtained by down-sampling the 2D high resolution sample at the finest scale with different down-sampling parameter ($d = 2^0, 2^1, 2^2, \ldots 2^n$). The down-sampling function is the same as the one defined in Eq. (3.5), except that all values between 0 and 1 are considered as 0.5 (grey). In other words, the level for each site at a coarse scale can be black (0), gray (0.5) or white (1), such that any site that is still black or white during down-sampling stays the same, while anything other than black or white is considered as gray (0.5).

6.3.2 Measurement Model

The measurement model describes how the information obtained from the measurement is included in the posterior energy function. In fact, the constraint J in Eq. (3.2) is considered as $J^k(Z^k; M)$, at each scale. The measurement, M, is a low resolution observation, and can be described by a forward model $f_m(\cdot)$ such that for Z at the finest scale

$$M = f_m(Z) \tag{6.8}$$

In the forward model f_m each measured pixel, m_I, in M is a representative of a set of d sites in Z, such that

$$m_I = \frac{1}{d} \sum_{j \in I} z_j^k \tag{6.9}$$

where z_j is a site in the binary random field Z at the finest scale. In other words, each m_I corresponds to the average of gray values of a set of pixels at the finest scale. Therefore, since we are considering a binary field at the finest scale (0 for black and 1 for white), each measured pixel corresponds to the fraction of *white* pixels in the corresponding set of pixels at the finest scale.

The forward model f_m is able to describe the relationship between the measurement and a simulated random filed Z at the *finest* scale. Therefore for the non-hierachical approaches, this type of forward model is sufficient in defining the measurement model in the posterior energy. However, for the hierarchical approaches the relationship between the measurement and a random field at an intermediate scale needs to be defined differently.

In contrast to the finest scale, it is not really possible to define such a forward model directly for the coarse scales. The relationship between the gray values and the measured pixels at the coarse scales is not pre-defined such as Eq. (6.9), rather it remains unclear. For example, a value 0.8 in the measurement means 80% of the pixels of the corresponding set of sites at the *finest* scale should be white, while at a given coarse scale, this value does not always mean that much white in the corresponding set of pixels, since some pixels may be gray at that scale. Therefore, we can not use the same forward model defined in Eq. (6.9) to describe the relationship between the measurement and a random field at a coarse scale. More specifically, the value 0.5 for a site at the coarse scale means a notion of uncertainty, and can not be counted as a real value in Eq. (6.8). According to the examples shown in Fig. 6.2, the value 0.5 at the coarse scales does not necessarily means half white and half black at the fine scale, e.g. we can see that some of the large white areas, even with small portion of black, correspond to gray at the coarse scales.

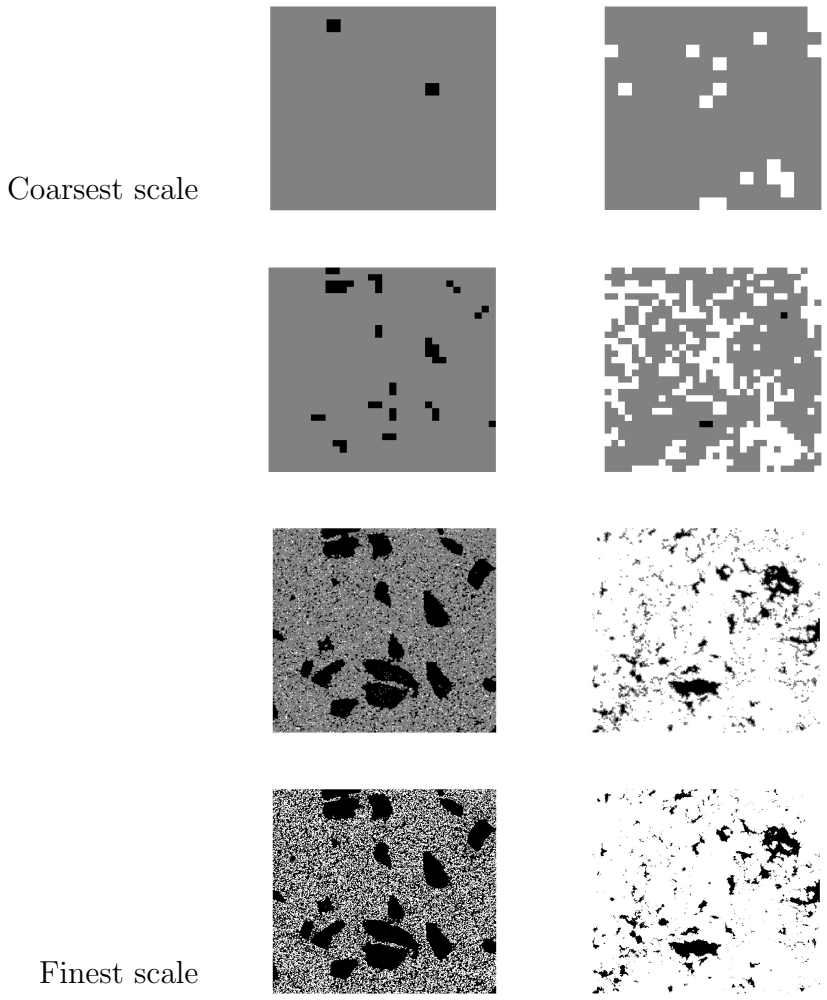


Figure 6.2: Two examples of porous media at different scales. On the left, most of the gray at the coarse scales contain almost equal number of black and white at the fine scale , except near large black area. However on the right, the coarse-scale gray corresponds to mostly white at the fine scale, since the fraction of white at the finest scale is relatively large. The gray values at the coarse scales do not necessarily corresponds to an even distribution of black and white at the fine scale.

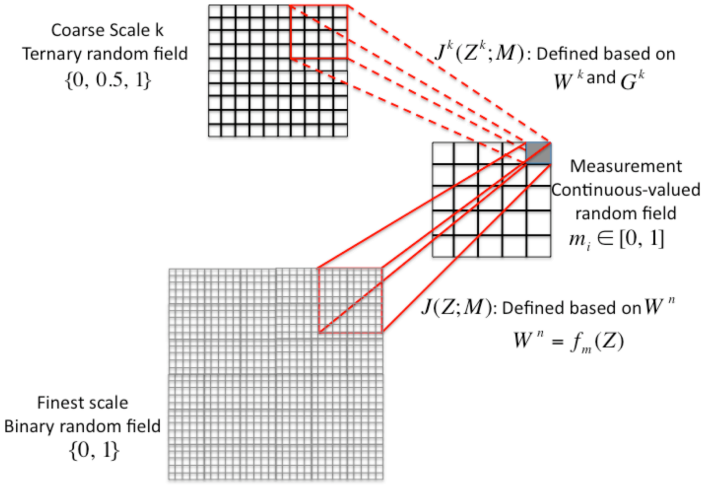


Figure 6.3: The relationship between the measurement and the fine and coarse scales.

The measurement model proposed here originates from studying the fraction of white and gray at each scale within a measured pixel. The fraction of gray (G^k) and white (W^k) at scale k are defined as

$$G^k = f_g^k(Z^k), \qquad W^k = f_w^k(Z^k). \tag{6.10}$$

The random fields G^k and W^k have the same resolution as the measurement M, and more specifically, a datum g_I in G^k (or w_I in W^k) is equal to the fraction of gray (or white) in the corresponding $2^k \times 2^k$ block of pixels in Z^k. It is obvious that at the finest scale (i.e. when $k = n$), $f_w^n(Z^n) = f_m(Z^n)$ and $f_g^n(Z^n) = 0$. Moreover, the measurement model at the finest scale is defined solely based on the fraction of white (W), while at a given coarse scale k, both fraction of gray (G^k) and fraction of white (W^k) are required to be considered in the measurement model, as shown in Fig. 6.3.

Since we do not have any gray at the finest scale, the constraint J^n at the finest scale is defined as

$$J^n(Z^n; M) = \|M - f_m(Z^n)\|. \tag{6.11}$$

To define the constraint term for the scales above the finest, we have studied the relationship between a measured pixel, m_I, and the corresponding g_Is, for several pixels, in a set of high resolution training data. Fig. 6.4 (a) shows this relationship as the scatter plot of g_I versus m_I at different scales. As can be seen from this

figure, there is a non-linear and complex relationship between these two variables. This relationship would be difficult to model.

On the other hand, according to the definition of the forward model f_m, given in Eq. (6.9), a given m_I in the measurement, in fact, shows the fraction of white, w_I in the corresponding set of pixels at the *finest* scale. Therefore, for an estimate at the finest scale, to obey the measurements, the residual $r_I = m_I - w_I$ (for all I) should be zero, while at the coarse scale r_I can be greater than zero, and consequently the more gray we have, the greater r_I will be. Thus, *we propose to consider the relationship between r_I and g_I*, shown in Fig. 6.4 (b), instead to define the measurement model at an intermediate scale. This new relationship does not contain the complexity shown in Fig. 6.4 (a). Thus we end up with a much simpler model, such that even at the scales close to the finest scale, it can be easily defined based on a simple linear regression.

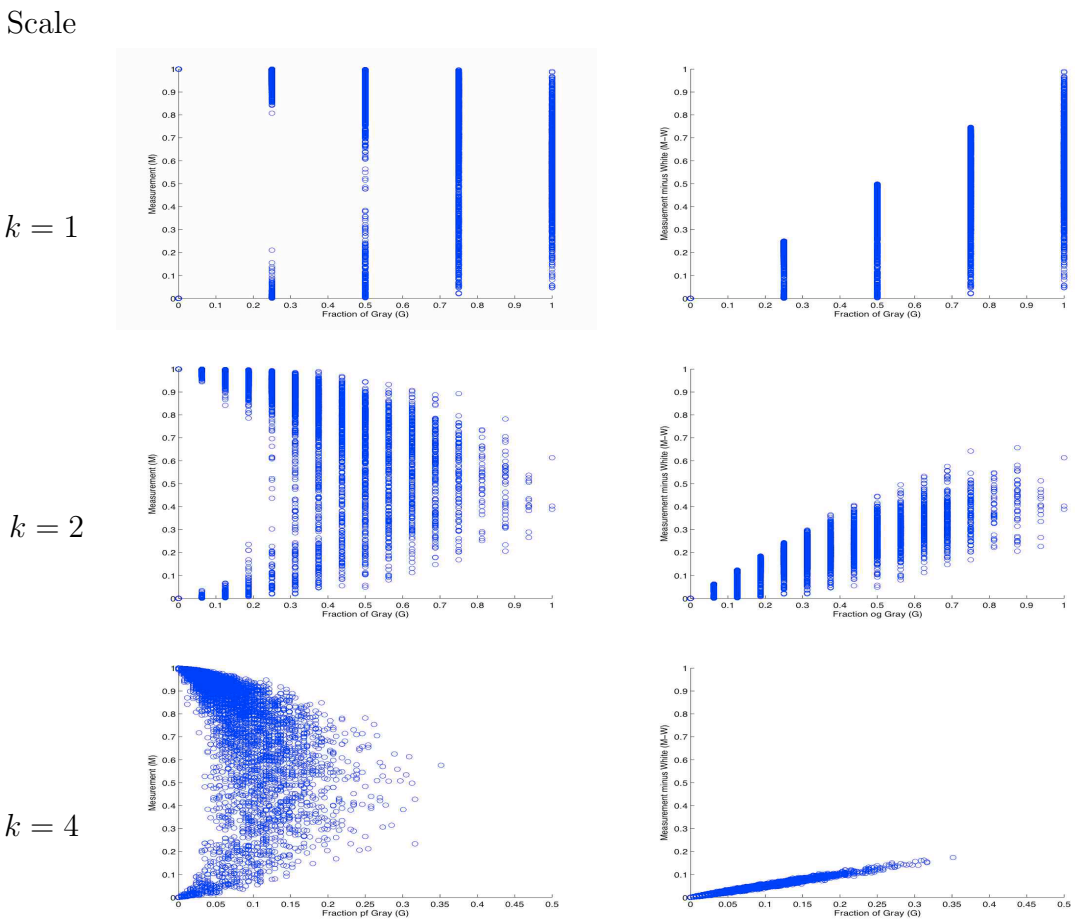
The results in Fig. 6.4 (b) are obtained based on multiple runs and different training samples. At the scales near the measured scale, the relationship between g_Is and r_Is gets complex, such that at the two coarsest scales below the measured scale, (when $k = 1$ and $k = 2$), it can not be described with a linear parametric model. For these scales we propose to consider a non-parametric model.

According to multiple runs for different types of porous media images, the following facts are observed when studying the relationship between r_I and g_I

- The relationship can be approximated by a linear parametric model for scales $k > 2$, as shown in Fig. 6.5
- For the two coarsest scales below the measured scale ($k = 1$ and $k = 2$), the relationship can not be represented by a parametric linear model. Fig. 6.6 showing the estimated linear model for scale $k = 1$ and $k = 2$ confirms this fact.
- In Fig. 6.4 (b) we see that the data are always on the lower triangle of the space, that is $r_I = m_I - w_I \leq g_I$. This inequality is always true and independent of the pattern of data, according to the following Lemma.

  Lemma. For every measured pixel m_I, and the corresponding fraction of white, w_I^k and fraction of gray, g_I^k, at an arbitrary scale k, we have $m_I - w_I^k \leq g_I^k$.

  Proof: Suppose that at scale k , the variables w_I^k and g_I^k correspond to the fraction of white and gray in a block of L pixels. Suppose that l_w pixels out



(a) Fraction of gray versus M (b) Fraction of gray versus $M - W$

Figure 6.4: (a) and (b) shows the scatter plot of fraction of gray versus the measurement and the fraction of gray versus "measurement minus fraction of white", respectively for the Carbonate Rock data shown in Fig. 6.2. Plots in (b) are simpler to be modeled than those in (a).

of the L pixels are white and l_g pixels are gray. Then we know that

$$w_I^k = \frac{l_w}{L}, \qquad g_I^k = \frac{l_g}{L}. \tag{6.12}$$

Each gray pixel, x_j^g, at scale k corresponds to a block of N pixels, (B_j), at the finest scale. For every gray pixel we attach a label wg_j which is the fraction of white in B_j. Therefore, a measured value m_I can be re-written as

$$m_I = \frac{l_w \ + \ \sum_{j=1}^{l_g} wg_j}{L} \tag{6.13}$$

Then according to Eq. (6.3.2), we have

$$m_I - w_I^k = \frac{\sum_{j=1}^{l_g} \ wg_j}{L} \tag{6.14}$$

Since $wg_j \leq 1$ then we have $\sum_{j=1}^{l_g} wg_j \leq l_g$, and therefore $m_I - w_I^k \leq g_I^k$. (Q.E.D.)

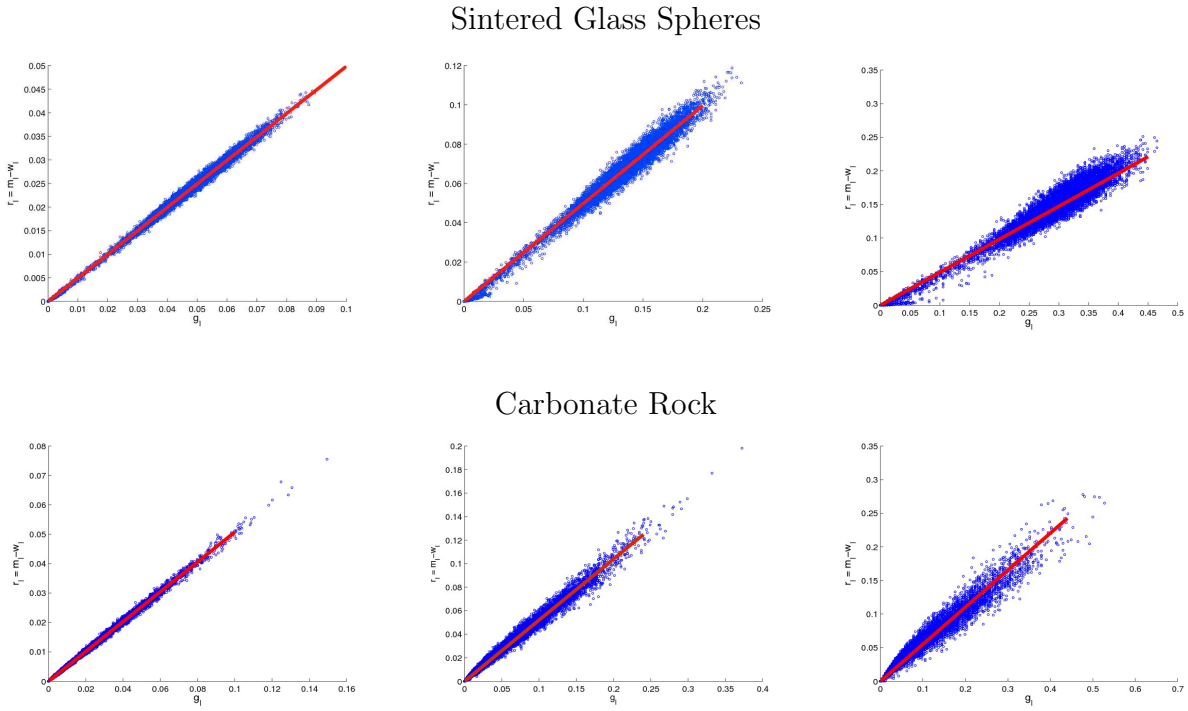
6.3.3 Linear parametric measurement model

The parametric model is defined when $k > 2$. This model is characterized with two parameters: a^k and v^k. The parameter a^k is the slope of the line obtained by a linear regression of g_I and r_I. The variance v^k represents the average deviation of data from the estimated linear model. These parameters are learned from the high resolution training data. Fig. 6.7 shows the estimated linear model for each scale $k > 2$. The measurement is considered to be nine scales coarser than the original image.

According to this model, the constraint term J^k in the hierarchical posterior energy function is defined as

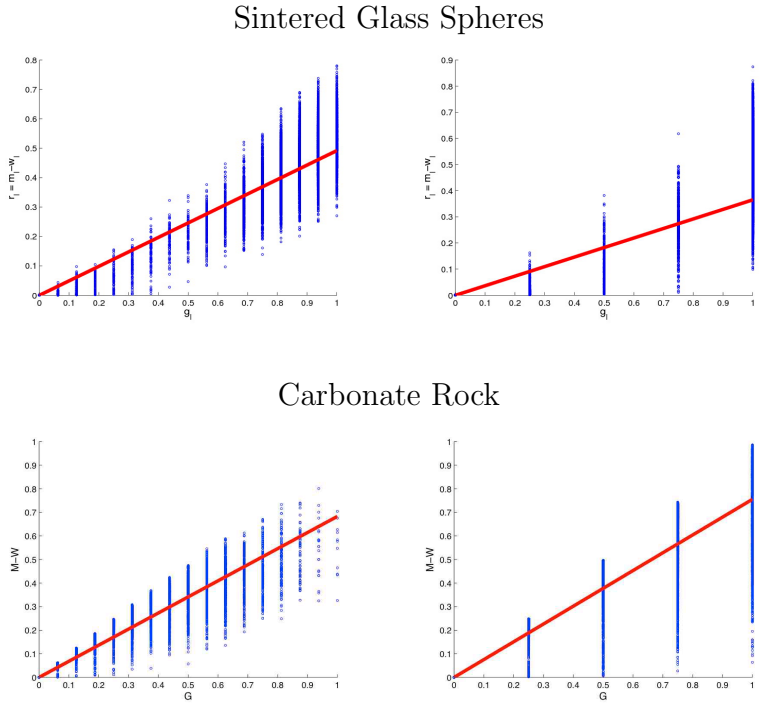
$$J^k(Z^k; M) = \frac{\|M - W^k - a^k \ G^k\|^2}{v^k + \epsilon}, \qquad when \ 2 < k < n \tag{6.15}$$

where $W^k = f_w^k(Z^k)$ and $G^k = f_g^k(Z^k)$ are the random fields corresponding to fraction of gray and white respectively, and ϵ is a small enough number to avoid dividing by zero in the model.



(a) $k = 6$ (b) $k = 5$ (c) $k = 4$

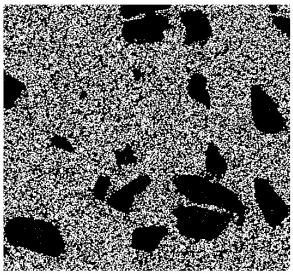
Figure 6.5: The relationship between fraction of gray , g_I and measurement-minus-white, r_I is estimated with a linear model (red line) for two different types of samples. The deviation from the linear model increases as we go coarser and coarser. The deviation from the line can also be considered as a single parameter in the model.

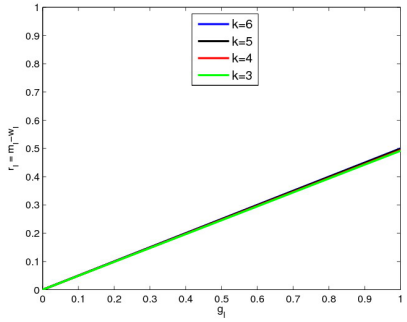


(a) $k = 2$ (b) $k = 1$

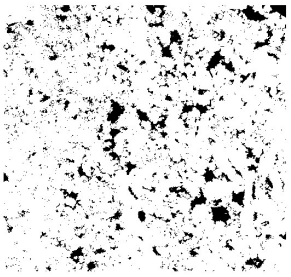
Figure 6.6: The relationship between fraction of g_I and r_I is estimated by a linear model (red line). The linear model can not represent the pattern of the data at the two coarsest scales, since the deviation from the linear model can not be defined by a single parameter.

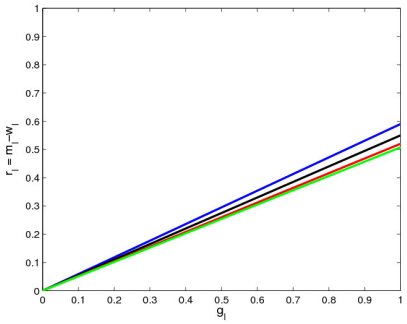
Sintered Glass Spheres





Carbonate Rock





(a) Training data (b) Estimated linear model

Figure 6.7: The linear parametric model for $k > 2$, for two different examples.The slopes are almost equal to 0.5 for all scales, which means that at a given intermediate scale, half of the pixels corresponding to gray in G should turn to white.

6.3.4 Non-parametric measurement model

The non-parametric model is defined for the two coarsest scales below the measured scale, i.e when $k = 1$ and $k = 2$, since at these scales the relations between g_I and r_I can not be modeled with a simple parametric model. The non-parametric model is defined based on the conditional probability of having a certain fraction of gray given a measured value in the measurement, or equivalently r_I. This conditional probability can be re-written according to Bays rule, as

$$P(G^k = G^*|M - W^k = R^*) \propto P(M - W^k = R^*|G^k = G^*)\ P(G^k = G^*). \quad (6.16)$$

where $k = 1,\ 2$. Then J^k is modeled as the logarithm of the conditional and marginal probabilities defined in (6.16), i.e.

$$J(Z^k; M) = |log(P(M - W^k = R^*|G^k = G^*))| + |log(P(G^k = G^*))| \quad (6.17)$$

where $W^k = f_w^k(Z^k)$ and $G^k = f_g^k(Z^k)$, based on the forward models defined in Eq. (6.10)

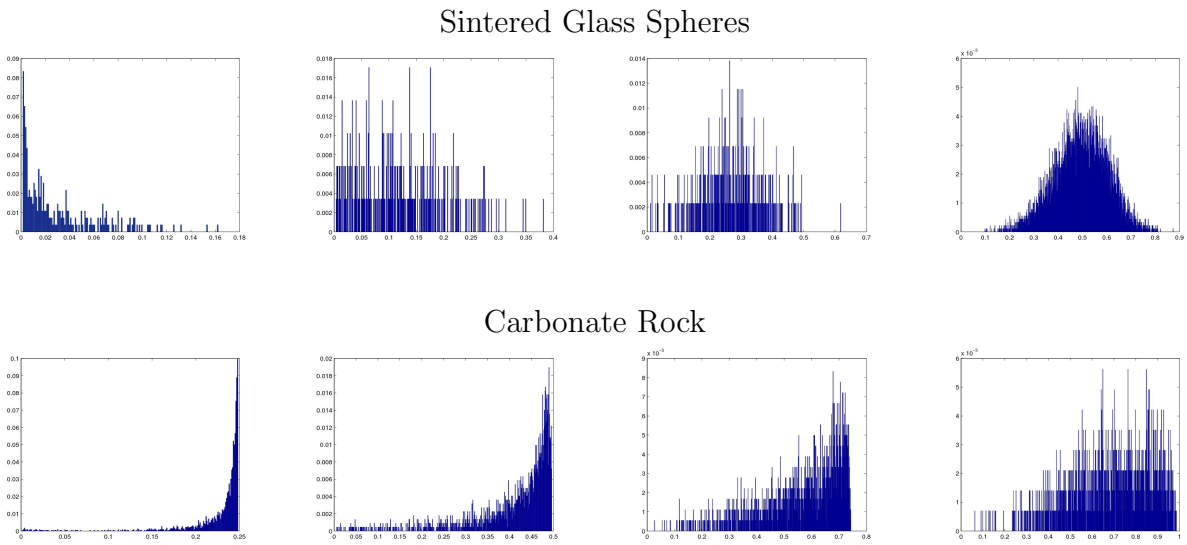
Non-parametric estimation of the conditional and marginal probabilities

The conditional probability in Eq. (6.16) can be estimated using a parametric or non-parametric probability density estimation, However, as can be seen from Fig. 6.8 in which the normalized histogram of r_I for different values of g_I are shown, the data do not follow a smooth shape of well-known probability distribution functions, therefore a non-parametric probability density estimation approach is proposed here.

The conditional and marginal probability in Eq. (6.16) can be estimated using non-parametric probability density function (pdf) methods, such as kernel density estimation [60, 61]. The kernel estimation methods are based on placing a window function $\mathcal{K}$, known as the kernel function, at the observations, and then summing over the kernels to form the probability density function. Given observations $x_i,\ i = 1, \ldots, n$, the probability density is defined

$$\hat{p}(x) = \frac{1}{Nh} \sum_{i=1}^{N} \mathcal{K}(\frac{x - x_i}{h}) \quad (6.18)$$

where h is the width of the kernel function, also called smoothing paramter or bandwidth. Rather than a fixed smoothing parameter, one can adjust h based on the



(a) $g_I = 0.25$ (b) $g_I = 0.5$ (c) $g_I = 0.75$ (d) $g_I = 1$

Figure 6.8: The normalized histogram of r_I at scale $k = 1$ for different fraction of gray, g_I, for the training data shown in Fig. 6.7.

density of the observations, leading to the *variable* kernel density estimation [61]. Considering $d_{j,s}$ to be the distance from x_j to the sth nearest point in the set comprising the other $N - 1$ data points, the variable kernel estimate is defined as

$$\hat{p}(x) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{h\ d_{j,s}} \mathcal{K}(\frac{x - x_i}{h\ d_{j,s}}) \tag{6.19}$$

where h is the overall smoothing parameter. The kernel width placed on point x_j is proportional to the density of data around that point, such that the data points in sparse regions have flatter kernels than the ones in dense regions. Choosing $s = \sqrt{n}$ has proven to be a wise choice [61].

To estimate the conditional probability in Eq. (6.16) variable kernel density estimation defined in Eq. (6.18)is considered. When $k = 1$, the variable g_I corresponds to a 2×2 block in Z^1, so the state space for g_I is the set $\{\frac{i}{4} | i = 0, 1, \ldots, 4\} = \{0,\ 0.25,\ 0.5,\ 0.75,\ 1\}$, and when $k = 2$, it is the set $\{\frac{i}{16} | i = 0, 1, \ldots, 16\}$. Therefore, we estimate five probabilities at scale $k = 1$ and 17 probabilities at scale $k = 2$, using variable kernel pdf estimation method. The estimated pdfs are also normalized to form a probability distribution function. Fig. 6.9 shows examples of the conditional probability distribution for different values of g_I at scale $k = 1$, and for two types of porous media examples, shown in Fig. 6.7 (a).

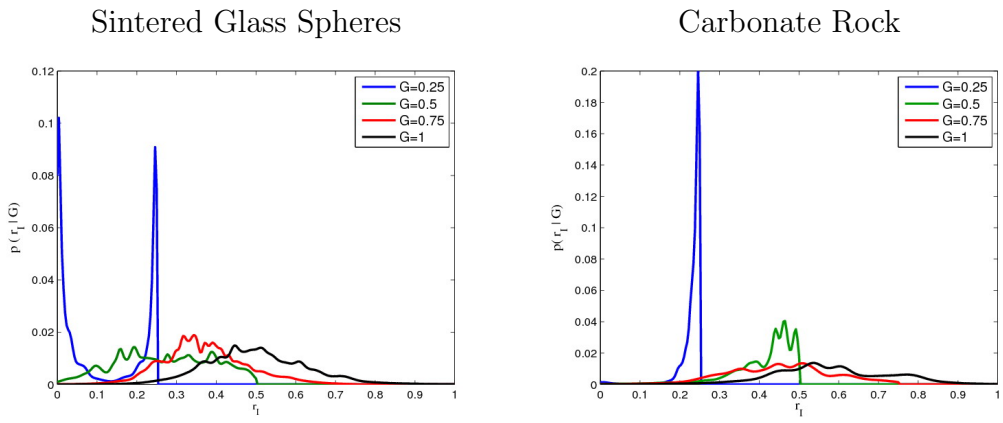


Figure 6.9: The conditional probability distribution of $M - W$ given a certain fraction of grey, in one scale below the measured scale, $k = 1$, for two different samples. The probabilities are learned non-parametrically from the high resolution training data in Fig. 6.7

The marginal probability of gray $P(G^k)$ is estimated using the histogram probability estimation method [61]. However, rather than considering bins with a specific width, the histogram is calculated at a given value of g_I, since g_I is a discrete variable. Fig. 6.10 shows the marginal probability distribution of gray at scale $k = 1, 2$, for two different images.

6.4 Hierarchical Sampling

The reconstruction task is done by Gibbs sampling from the posterior probability distribution [34]. The hierarchical sampling used here is a top-down approach based on Gibbs sampler along with simulated annealing, as proposed in [6, 18, 42]. We initialize at the measurement scale, $k = 0$, based on the measurement such that every pixel in Z^0 stays white/black if the measurement is white/black otherwise it is considered as gray. For the noisy measurement we can start with purely gray Z^0. Then Z^0 is projected to the next scale, $k = 1$. At this scale the black or white pixels in the projected random field are frozen while we anneal only over the gray pixels, using the flat annealing approach proposed in Chapter 3 and 4. This procedure continues until the finest scale is reached. The sample generated at each scale is consistent with the prior model at that scale and the measurement. Fig. 6.11 shows the reconstruction based on the proposed modeling and sampling approach. The reconstructed image and the original image does not need to be exactly the same,

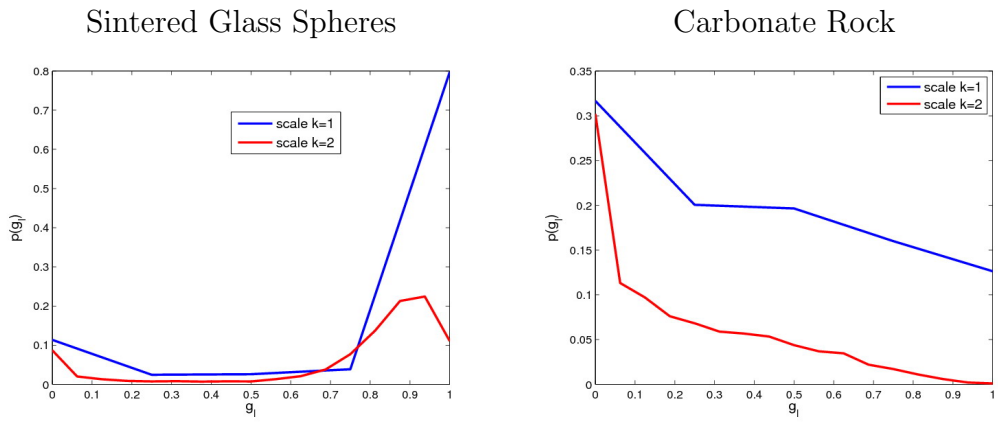


Figure 6.10: The estimated marginal probability distribution of the fraction of gray, G, at scale $k = 1$, for two different samples. The probabilities are learned non-parametrically from the high resolution training data in Fig. 6.7. The image used to generate (a) consists of small black/white structures, therefore large fraction of gray ($g_I > 0.5$) are more probable to happen than the small ones at the coarse scales, while the image used to generate (b) has large areas of white and black and even at a very coarse scale we do not expect to see lots of gray.

since we are generating random samples from the model, rather they need to obey the same statistics.

In order to evaluate the performance of the proposed measurement model, we have compared the proposed model with another model, we call it as average model, in which the gray value 0.5 is directly considered in the measurement model, such that

$$J^k(Z^k; M) = \|M - f_m^k(Z^k)\| \tag{6.20}$$

where f_m^k is similar to the down-sampling forward model defined in Eq. (6.8) and Eq. (6.9), but with different down-sampling parameter, such that $d = 2^k$. The comparison results are shown in Fig. 6.12. Since the measurement values are very close to zero in Fig. 6.12 the measurement energy defined based on Eq. (6.20) likes to change most of the pixels to black. Therefore, the results have insufficient white areas. Comparing the results in (a) with (b), we observe that the proposed measurement model is able to provide more *accurate* assertion for the fraction of black, gray and white at each scale.

The focus of this chapter is on how to incorporate the continuous-valued measurement with a discrete-valued unknown at an intermediate coarse scale in the hierarchical model, and more specifically how to define a posterior energy func-

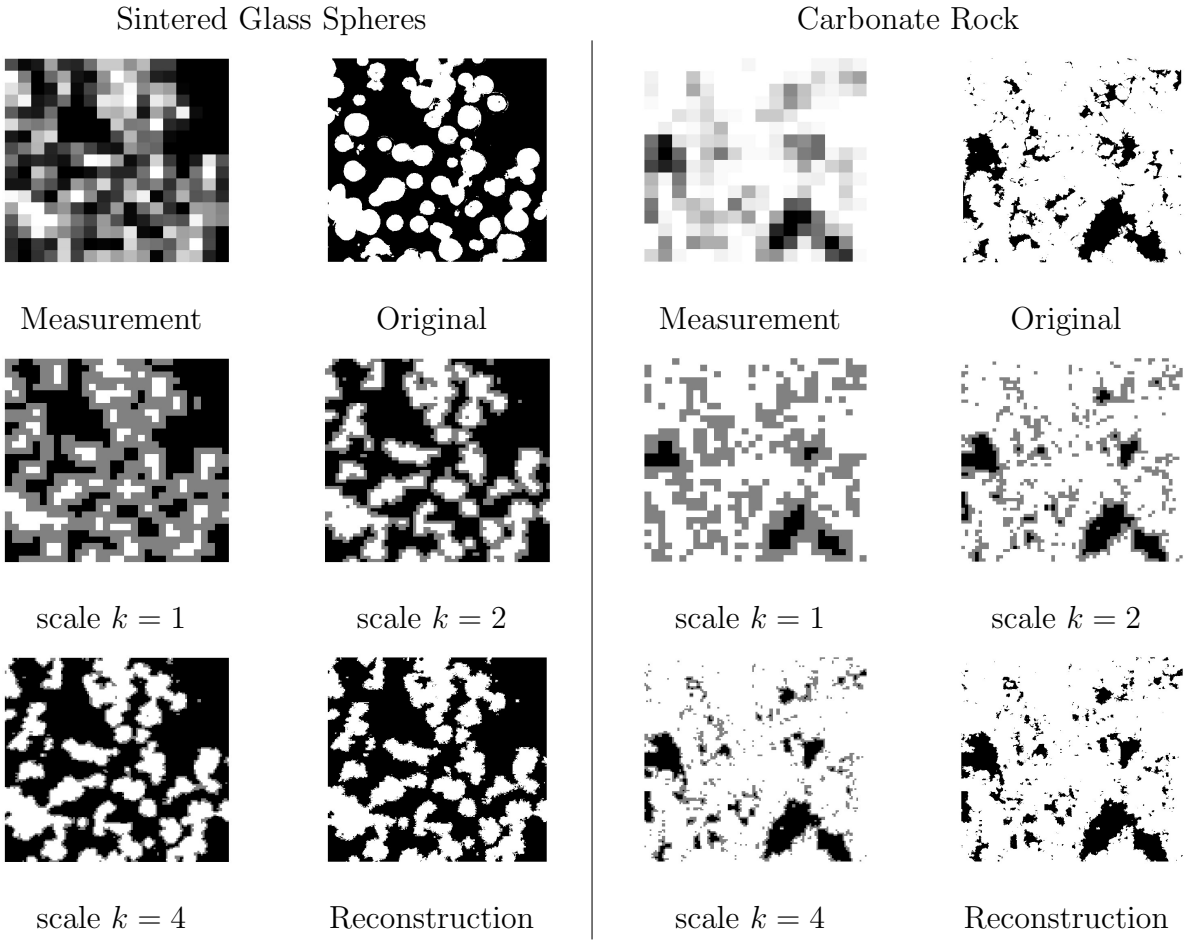
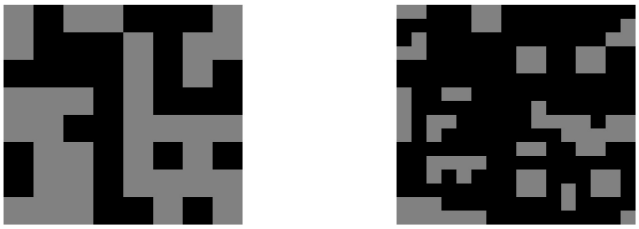


Figure 6.11: Reconstruction of two different porous media samples at different scales using the proposed hierarchical approach.

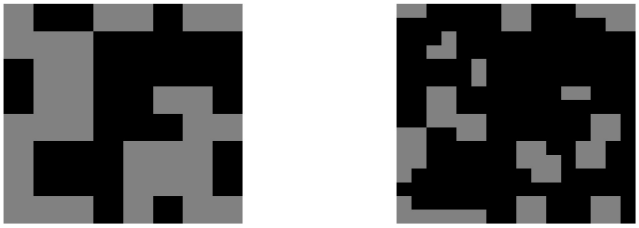
(a) Measurement

scale $k = 1$ scale $k = 2$



(b) Original data at the intermediate scales

(c) Reconstruction using average measurement model



(d) Reconstruction using proposed measurement model

Figure 6.12: Comparing the performance of the proposed model with a simple averaging model described in Eq. (6.20). The simple averaging model does not have any information on the fraction of gray at each scale and treats the gray pixels as 0.5 in the measurement model, leading to large fraction of black in the reconstructed samples.

tion at each level of the hierarchy. The studies in [6] and [18] show a significant improvement in the time required to generate large reconstruction using hierarchical approaches. One can study and evaluate the performance of the proposed model in terms of different aspects, such as the resolution of the measurement, the consistency with the training data in terms of statistical features, and also the computational time. However, this chapter's emphasis is on how to define a measurement model for the hierarchical approach, and such evaluations can also be visited in future research directions.

6.5 Summary

Reconstructing binary random images of porous media becomes intractable when the size of the images increases. Due to the computational complexity of the non-hierachical porous media reconstruction approaches, we have proposed a hierarchical posterior model in which the continuous-value measurement is related with a discrete-value unknown at each scale. Based on the relationship between the measurement and the fraction of grey at each scale, we define a parametric and non-parametric model. The parametric measurement model is used for the fine scales where the relationship can be described with a linear model. The non-parametric measurement model is defined for the coarse scales, where the relationship is difficult to be described with a simple linear model. Comparing with a simple averaging model, the proposed model is more powerful in generating meaningful strautures at each scale [62].

Chapter 7

Conclusions and Future Directions

This chapter contains a summary of research contributions developed in the thesis and also potential research directions for future work.

7.1 Summary and Conclusion

In this thesis a statistical data fusion framework is proposed in order to fuse the information obtained from porous media images at different resolutions. Based on the proposed framework, a constrained sampling approach is introduced for porous media reconstruction purposes. Despite the current estimation /optimization methods in super-resolution, image registration and multi-resolution analysis, the proposed approach is developed for fusing image data obtained from the same porous material but different scenes. Based on the proposed approach typical samples of porous media are generated which are consistent with the measurement and the prior model. The main contributions and findings of the thesis are as follows,

- Statistical fusion of different types of measurements with the prior model for two-scale porous media reconstruction

  In two-scale porous media reconstruction, only the large scale structures are resolved by the low resolution measurement. In this thesis a statistical data fusion approach is proposed based on posterior sampling for two-scale porous media reconstruction. In the proposed approach the statistical model learned from the high resolution data is fused with the measurements to construct a posterior model. We have considered two types of measurement: low resolution local porosity and surface-to-volume ratio (S/V) measurement. Since the

low resolution local porosity measurement can only resolve large scale structures, we propose to add surface-to-volume ratio measurement in the model as well. The S/V measurement provides information on the size of structures at different scales. The proposed statistical fusion approach provides samples which are more correlated and consistent with the real porous media samples, as compared to the similar methods in the literature.

- Constrained sampling using simulated annealing

  The stochastic reconstruction of porous media implies using MCMC methods along with simulated annealing to generate samples from the prior model. However porous media encompass variable structures at different scales, and annealing over a single model without any particular constraints does not provide samples with desired variability of the structures at different scales. We have proposed a constrained sampling approach by which an additional information is added into the prior model as the constraint. The proposed constrained sampling approach allows annealing down to a finite temperature while generating more reliable structures, as compared to annealing down to a zero temperature. The variability of the samples generated by the proposed method is greater than the unconstrained annealing down to zero temperature. The proposed method is able to generate samples which are more consistent with porous media images, comparing with similar methods in the literature.

- Hierarchical posterior sampling of porous media images

  Hierarchical methods in image analysis are able to decrease the computational complexity, specifically caused by simulated annealing and MCMC approaches. The research interest in hierarchical sampling of binary images [6, 18] is limited to the case where the measurements are not involved. Based on the hierarchical ternary representation proposed in [18], we have proposed a hierarchical posterior model in which the measurement is related with an unknown at each scale. A new measurement model is proposed for each scale of the hierarchy based on the relationship between the continuous-value measurement and the discrete-value ternary random field at that scale. Comparing the result with the case in which a simple averaging measurement model is defined, we have observed that the proposed measurement model provides reliable and valuable information on the samples at the coarse scale.

7.2 Future Research Directions

The proposed approach in this thesis provides a general framework for fusing different types of information obtained from porous media images. A variety of promising results obtained based on the proposed approach suggest future research directions in 3D reconstruction of porous media. In addition to the 3D reconstruction, annealing with constraint can be re-visited again in order to derive other types of criteria to study the variability of the samples. Also the method can be extended to include the measurement as well. Collaborative work with other colleagues in using hierarchical methods in porous media reconstruction [63] provides possible research directions in extending the hierarchical methods to more complicated prior models and also considering other types of measurements in the model. Moreover, the simultaneous annealing strategy over multiple scales in the hierarchical framework is a new research idea which requires an extensive study and development. Possible future research directions are as follows:

- 3D reconstruction of porous media is a crucial step in studying different properties of porous materials. Based on the proposed framework for data fusion, the reconstruction task can be generalized into 3D space. However, defining a 3D prior model and relating the model with the 3D samples requires more elaborations and is a potential future research opportunity.

- Annealing with constraints can be generalized into more complex situations in which the measurements and other types of prior models are also involved. In such cases, defining other types of constraints in the energy function and choosing a suitable annealing strategy to satisfy different constraints is still an open problem and requires more elaborations. Moreover, more sophisticated criteria than standard deviation can be developed to study the variability of the samples.

- The hierarchical model can be generalized to handle complex non-stationary models as well as other types of measurements such as surface-to-volume measurement. Involving surface-to-volume ratio measurement requires defining different measurement models, which can be considered as a future research direction.

- The simultaneous annealing strategy in the hierarchical framework, proposed very briefly in the thesis is a new research area which can influence the optimization strategies used in the hierarchical frameworks significantly.

Bibliography

[1] P.M. Adler. *Porous Media, Geometry and Transports*. Butterworth-Heinemann series in chemical engineering. Butterworth-Heinemann, 1992.

[2] K. Sobczyk and D.J. Kirkner. *Stochastic Modeling of Microstructures*. Modeling and simulations in science, engineering and technology. Birkhauser, 2001.

[3] S. Torquato. *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*. Springer-Verlag, 2002.

[4] G. S. Padhy, C. Lemaire, E. S. Amirtharaj, and M. A. Ioannidis. Pore size distribution in multiscale porous media as revealed by DDIF-NMR, mercury porosimetry and statistical image analysis. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 300(1):222–234, 2007.

[5] H. Okabe and M. J. Blunt. Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics. *Water Resource Research*, 43, 2007.

[6] S.K. Alexander, P. Fieguth, M.A. Ioannidis, and E.R. Vrscay. Hierarchical annealing for synthesis of binary images. *Mathematical Geosciences*, 41(4):357–378, 2008.

[7] P. Fieguth. Hierarchical posterior sampling for images and random fields. *IEEE, ICIP*, 2003.

[8] C. L. Y. Yeong and S. Torquato. Reconstructing porous media. *Physical Review E*, 57(1):495–506, 1998.

[9] C. L. Y. Yeong and S. Torquato. Reconstructing random media ii. three-dimensional media from two-dimensional cuts. *Physical Review E*, 58(1):224–233, 1998.

[10] C. Manwart, S. Torquato, and R. Hilfer. Stochastic reconstruction of sandstones. *Physical Review E*, 62(1):893–899, 2000.

[11] M.A. Ioannidis and I. Chatzis. On the geometry and topology of 3d stochastic porous media. *Journal of Colloid Interface Science*, 229:323–334, 2000.

[12] A. Mohebi and P. Fieguth. Posterior sampling of scientific images. *LNCS*, 4141:365–376, 2006.

[13] R. Chellappa and A. Jain. *MRF Theory and Applications*. Academic Press, 1993.

[14] D. Rajan and S. Chaudhuri. *Super Resolution Imaging*, chapter Super resolution imaging using blur as a cue. Kluwer, 2001.

[15] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li. A comparative analysis of image fusion methods. *IEEE Trans. on Geoscience and Remote Sensing*, 43(6):1391–1402, 2005.

[16] G. Piella. A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion*, 34:259280, 2003.

[17] Z. Liang, M.A. Ioannidis, and I. Chatzis. Geometric and topological analysis of three-dimensional porous media: Pore space partitioning based on morphological skeletonization. *Journal of Colloid Interface Science*, 221(24):13–24, 2000.

[18] W. R. Campaigne, P. Fieguth, and S. K. Alexander. Frozen-state hierarchical annealing. In *ICIAR*, pages 41–52, 2006.

[19] M. Bertero and P. Boccacci. *Introduction to Inverse problems in Imaging*. CRC press, 1998.

[20] R. C. Aster, B. Borchers, and C. Thurber, editors. *Parameter Estimation and Inverse Problems*. Elsevier Academic Press, 2004.

[21] P. Fieguth. Statistical processing of multidimensional data. Lecture notes, 2001.

[22] M. Bertero and P. Boccacci, editors. *Introduction to Inverse Problems in Imaging*. CRC Press, 1998.

[23] G. E. P. Box and G. C. Tiao, editors. *Bayesian Inference in Statistical Analysis*. Wiley, 1992.

[24] M. Joshi. *A Class of Stochastic Models for Porous Materials*. PhD thesis, University of Kansas, Lawrence, 1974.

[25] X. Zhao, J. Yao, and Y. Yi. A new stochastic method of reconstructiong porous media. *Transport in Porous Media*, 69(1):1–11, 2007.

[26] P. Fieguth and J. Zhang. *Handbook of Video and Image Processing*, chapter Random Field Models. Academic Press, 2000.

[27] R. C. Gonzalez. *Digital Image Processing*. Pearsons Education, second edition, 2002.

[28] R. Chellappa and A. Jain. *MRF theory and applications*, chapter Image Modeling during 1980s: A Brief Overview. Academic Press, 1993.

[29] D. Geman. *Random Fields and Inverse Problems in Imaging*, volume 1427 of *Lecture Notes in Mathematics*. Springer-Verlag, 1990.

[30] S. T. Li, editor. *Markov Random Field Modeling in Computer Vision*. Springer, 1995.

[31] G. Winkler. *Image Analysis, Random Fields, and Markov Chain Monte Calro Methods*. Springer-Verlag, second edition, 2003.

[32] B. Chalmond. *Modelling and Inverse Problems in Image Analysis*, volume 155 of *Applied mathematical sciences*. Springer-Verlag, 2003.

[33] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer-Verlag, second edition, 2004.

[34] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and the bayesian restoration of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 6(6), 1984.

[35] A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Sciences*, 7(4):457–472, 1992.

[36] A. E. Rafery and S. M. Lewis. *Practical Monte Carlo*, chapter The number of iterations, Convergence Diagnostics and Generic Metropolis Algorithms, pages 115–130. Chapman and Hall, 1995.

[37] N. Metropolis, A.Rosenbluth, M.Rosenbluth, A. Teller, and E.Teller. Equation of state calculation by fast computing machines. *J. of Chemical Physics*, 21:1087–1092, 1953.

[38] J.C.Spall. Estimation via markov chain monte carlo. *IEEE Control Systems Magazine*, 23(2):34–45, 2003.

[39] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[40] D.Bertsimas and J.Tsitsiklis. Simulated annealing. *Statistical Science*, 8(1):10–15, 1993.

[41] A. Chardin and P. Prez. Semi-iterative inferences with hierarchical energy-based models for image analysis. In *Proc. International Workshop EMM-CVPR99: Energy Minimisation Methods in Computer Vision and Pattern Recognition*, pages 165–4. Springer LNCS, 1999.

[42] C. Graffigne, F. Heitz, P. Perez, F. J. Preteux, M. Sigelle, and J. B. Zerubia. Hierarchical markov random field models applied to image analysis: A review. volume 2568, pages 2–17. SPIE, 1995.

[43] J. Goodman and A. D. Sokal. Multigrid monte carlo method. conceptual foundations. *Physical Review D*, 40(6):2035–2071, 1989.

[44] B. Gidas. A renormalization group approach to image processing problems. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(2):164–180, 1989.

[45] F. Heitz, P. Perez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. In *Comput. Vision Graphics and Image Process*, pages 125–134, 1994.

[46] S. Stapf and S. Han, editors. *NMR Imaging in Chemical Engineering*. Wiley-VCH, 2006.

[47] A. E. Pomerants, E. E. Sigmud, and Y. Q. Song. Spatial heterogeneity length scale in carbonate rock. *Applied Magnetic Resonance*, 32:221–231, 2007.

[48] J.P. Butler, R. W. Mair, D. Hoffmann, M. I. Hrovat, R. A. Rogers, G. P. Topulos, R. L. Walsworth, and S. Patz. Measuring surface-area-to-volume ratios in soft porous materials using laser-polarized xenon interphase exchange nuclear magnetic resonance. *J. of Physics*, pages 297–304, 2002.

[49] S. Chaudhuri. *Super Resolution Imaging*. Kluwer, 2001.

[50] A. Mohebi and P. Fieguth. Constrained sampling using simulated annealing. *LNCS*, 2007.

[51] D. Geman. *Random Fields and Inverse Problems in Imaging*, volume 1427 of *Lecture Notes in Mathematics*. Springer-Verlag, 1990.

[52] A. Mohebi, P. Fieguth, and M. A. Ioannidis. Statistical fusion of two-scale images of porous media. *Advances in Water Resources, accepted*, 2009.

[53] P. P. Mitra, N. P. Sen, and M. L. Schwartz. Short-time behavior of the diffusion coefficient as a geometrical probe of porous media. *Phys. Rev. B*, 47(14):8565–8574, 1993.

[54] Y. Liu and P. Fieguth. Image resolution enhancement with hierarchical hidden fields. *Lecture Notes in Computer Science*, 5627:73–82, 2009.

[55] A. Mohebi and P. Fieguth. Statistical fusion and sampling of scientific images. In *ICIP*, pages 1312–1315, 2008.

[56] A. Mohebi, P. Fieguth, and M. A. Ioannidis. Modeling and reconstruction of two-scale porous media using mri measurements. In *4th Biot conference on Poromechanics*, 2009.

[57] K. J. Falconer. *Techniques in Fractal Geometry*. Wiley, 1997.

[58] A. Mohebi and P. Fieguth. Constrained sampling using simulated annealing. In *ICIAR*, pages 198–209, 2007.

[59] H. Szu and R. Hartley. Fast simulated annealing. *Physic Letters A*, (122), 1987.

[60] R. O. Duda, P.E. Hart, and D. G. Stork, editors. *Pattern Classification*. Wiley, second edition, 2000.

[61] W. Silverman, editor. *Density Estimation*. Chapman and Hall, 1986.

[62] A. Mohebi, Y. Liu, and P. Fieguth. Hierarchical sampling with constraints. *Lecture Notes in Computer Science*, 5627:23–32, 2009.

[63] Y Liu, A. Mohebi, and P. Fieguth. Modeling of multi-scale porous media using multiple markov random fields. In *4th Biot conference on Poromechanics*, 2009.

Appendix A

Matlab Code for Computing the Correlation

This MATLAB code, written by Dr. Paul Fieguth, computes the correlation between an original binary data and also a set of reconstructed samples are generated across different scales.

```
% the passed matrices f is original
% matrices m1, m2, m3, ... are compared against
% all matrices are of the same size and binary

% numruns indicates how many times the code is to be run, providing
% error bars on the returned values

% the 'per' flag is for periodic fields, which are then
% randomly shifted for better averaging; set to zero for non-periodic

% 'maxscale' indicates the greatest scale at which to test; points within
% maxscale of the boundary for non-periodic images will not be used

% scalestep controls fineness of scale sampling

% numpass allows a specifying of number of random shifts in learning scales
% set to zero or negative or [] to get default

% returned matrix fn gives an image of the learned scales in f
% returned array se gives error std dev, only if numruns > 1
```

```
% Paul Fieguth
% February, 2008


function [sr,sc,fn,se] = scalecorr( numruns, f, per, maxscale, scalestep, ...
numpass, m1, m2 ,m3, m4, m5, m6, m7, m8 )

if (numruns>1),
  mstr = '';
  for m=1:nargin-6, mstr = [mstr ', m' int2str(m)]; end;
  for i=1:numruns,
    eval( ['[sr,sc,fn] = scalecorr( 1 , f, per, maxscale, scalestep,...
      numpass' mstr ');'] );
num2str(sc)
    if (i==1),
      scs = sc; scss = sc.^2;
    else,
      scs = scs + sc; scss = scss + sc.^2;
    end;
  end;
  sc = scs/numruns;
  se = sqrt( scss/numruns - sc.^2 );
  return
end;

sr = 1:scalestep:(1+floor(log(maxscale)/log(2)));
nonpersiz = floor((size(f)-maxscale) ./ (2^(max(sr)-1)))*(2^(max(sr)-1));
nonpershift = size(f)-nonpersiz;
if length(numpass)>0,
  if numpass<=0, numpass = []; end;
end;
if length(numpass)==0,
  numpass = round(sqrt(max(size(f))));
end;

sc = 0*sr;
si = 0*sr;
sr = [sr inf];

fsave = f;
```

```
for m=1:nargin-6,
f = fsave;
eval(['g = m' int2str(m) ';']);

% get both f and g to +/- 1 values
fl = find(f<min(f(:))+0.5); fh = find(f>min(f(:))+0.5);
f(fl) = -1; f(fh) = 1;
gl = find(g<min(g(:))+0.5); gh = find(g>min(g(:))+0.5);
g(gl) = -1; g(gh) = 1;

fn = 0*f;
if (per > 0),
  % compute scale over random rotations if periodic
  for i=1:numpass,
    ro=1+floor(size(fn,2)*rand);
    co=1+floor(size(fn,2)*rand);
    f=[f(:,(1+ro):end) f(:,1:ro)];
    f=[f((1+co):end,:); f(1:co,:)];
    g=[g(:,(1+ro):end) g(:,1:ro)];
    g=[g((1+co):end,:); g(1:co,:)];
    fn=[fn(:,(1+ro):end) fn(:,1:ro)];
    fn=[fn((1+co):end,:); fn(1:co,:)];
    a=f;
    okmaskp = 0*f+1;
    okmaskn = 0*f+1;
    kronmask = 1;
    while (min(size(a))>2 & rem(size(a),2)==0),
      fn = fn + okmaskp .* (f>0.5) + okmaskn .* (f<-0.5);
      a = downsamp(a,[1 2],1.5);
      kronmask = kron( kronmask, ones(2) );
      okmaskp = okmaskp .* kron( a>0.5,kronmask );
      okmaskn = okmaskn .* kron( a<-0.5,kronmask );
    end;
  end;
  fn = fn / numpass;
else,
  % compute scale over random shifts if not periodic
  fc = 0*f;
  for i=1:numpass,
    ro = 1+floor(nonpershift(1)*rand);
    co = 1+floor(nonpershift(2)*rand);
    rr = ro:(ro+nonpersiz(1)-1);
```

```
    cr = co:(co+nonpersiz(2)-1);
    ft = f(rr,cr);
    fc(rr,cr) = fc(rr,cr) + 1;
    a = ft;
    okmaskp = 0*a+1;
    okmaskn = 0*a+1;
    kronmask = 1;
    while (min(size(a))>2 & rem(size(a),2)==0),
      fn(rr,cr) = fn(rr,cr) + okmaskp .* (ft>0.5) + okmaskn .* (ft<-0.5);
      a = downsamp(a,[1 2],1.5);
      kronmask = kron( kronmask, ones(2) );
      okmaskp = okmaskp .* kron( a>0.5,kronmask );
      okmaskn = okmaskn .* kron( a<-0.5,kronmask );
    end;
  end;
  fcg = find(fc>=numpass);
  fn(fcg) = fn(fcg) ./ fc(fcg);
  fn(find(fc<numpass)) = 0;
end;

for s=1:length(sr)-1,
  i = find(fn>=sr(s) & fn<sr(s+1));
  if length(i)>0,
    sc(s) = sc(s)+mean(f(i).*g(i));
    si(s) = si(s)+1;
  end;
end;
end;

sc = sc ./ si;
sr = sr(1:(end-1)) - 1;
```