Procedia Computer Science

Volume 51, 2015, Pages 1188–1197

ICCS 2015 International Conference On Computational Science

Statistical Inversion of Absolute Permeability in Single-Phase Darcy Flow

Thilo Strauss[1], Xiaolin Fan[2], Shuyu Sun[2], and Taufiquar Khan[1]*

[1] Department of Mathematical Sciences, Clemson University, Clemson, USA
[2] Division of Physical Sciences and Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal, Kingdom of Saudi Arabia

Abstract
In this paper, we formulate the permeability inverse problem in the Bayesian framework using total variation (TV) and ℓ_p $(0 < p \leq 2)$ regularization prior. We use the Markov Chain Monte Carlo (MCMC) method for sampling the posterior distribution to solve the ill-posed inverse problem. We present simulations to estimate the distribution for each pixel for the image reconstruction of the absolute permeability.

Keywords: Statistical inversion, Single-phase Darcy flow, Markov chain Monte Carlo method, Metropolis-Hastings

1 Introduction

Subsurface flow models in oil reservoir, underground aquifers, carbon dioxide sequestration are broadly employed to predict the fluid flow in the porous media [20, 5, 17, 16, 18, 19, 7]. Combination of Darcy's law and continuity equation is mostly used to describe the physics and dynamics of fluid in the porous media. Within these equations, several parameters should be given before running the model to obtain the solution. The most important one among them is the absolute permeability characterizing the penetrating ability for fluid in the porous media. It is not realistic to measure the value of the absolute permeability in the interior of the porous media of interest. Therefore researchers use inverse methods to estimate the value of the absolute permeability from the boundary measurements. There exists numerous approaches to solve the inverse problem for the single-phase Darcy flow for example using Levenberg-Marquardt [12], iterative Gauss-Newton [1] etc. The above mentioned studies focus on deterministic inversion. An alternative approach, the statistical inversion methods for example in electrical impedance tomography [14], can shed interesting insights into the reconstructions. In [2] some results on the uncertainty quantification of MCMC based image reconstruction has been investigated. Statistically, the sparsity regularization amounts to enforcing ℓ_p prior on the expansion

*Corresponding Author

 Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2015

doi:10.1016/j.procs.2015.05.291

coefficients in a certain basis similar to the deterministic approach [13]. In this paper, we investigate a statistical approach where we consider both TV and ℓ_p $(1 \leq p \leq 2)$ prior as well as non-convex ℓ_p $(0 < p < 1)$ priors. Furthermore, we use a fast algorithm to compute the solution using a special type of a Metropolis Hasting algorithm.

2 Mathematical Model

The modeling mathematical system consists of continuity equations and Darcy's law, and boundary conditions.

2.1 Darcy's law for incompressible single-phase flow in porous media

Let $\Omega \subseteq R^2$ be an open and bounded set with C^1 smooth boundary denoted by $\partial\Omega$. Darcy's law relates the Darcy velocity to the driven force including the pressure and gravity by

$$\mathbf{u} = -K\nabla p, K = \frac{kg}{\mu} = \frac{kg\rho}{\eta} \text{ in } \Omega \tag{1}$$

where $\mathbf{u}$, k, μ, ρ, g, η, represents the Darcy velocity, kinematic viscosity of fluid, density of fluid, gravitational constant and dynamic viscosity of fluid, respectively, and K represents the absolute permeability. In this paper, we formulate a statistical inverse problem to estimate the absolute permeability K.

2.2 Mass conservation for incompressible single-phase flow in porous media

The fluid is assumed to flow continuously, which leads to the mass conservation of fluid and is described by the following mass conservation equation

$$\frac{\partial\rho\phi}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = q_m \text{ in } \Omega \tag{2}$$

where q_m is mass source term (is positive when the media is injected into fluid) or sink term (is negative when the fluid is extracted out from the media). In this work, we assume the fluid to be incompressible here. As a result, Equation (2) leads to the following volumetric conservation equation

$$\nabla\cdot\mathbf{u} = q \text{ in } \Omega \tag{3}$$

where $q = \frac{q_m}{\rho}$ and $\partial\Omega = \Gamma = \Gamma_D \bigcup \Gamma_N$ (Γ_D,Γ_N stand for the Dirichlet boundary and Neumann portion of the boundary respectively). The boundary conditions are given as follows:

$$p = p_B \text{ on } \Gamma_D \tag{4}$$

$$\mathbf{u}\cdot\mathbf{n} = u_B, \text{ on } \Gamma_N \tag{5}$$

where p_B and u_B are the pressure on Dirichlet boundary and the normal component of velocity on Neumann boundary respectively, and $\mathbf{n}$ is the outward normal unit vector towards on boundary.

In the forward model, the unknowns p, $\mathbf{u}$ can be solved using Equation (1) - Equation (5) given the permeability, viscosity, density. In the subsequent section we provide a numerical algorithm to solve the model.

3 Mixed Finite Element Methods For The Forward Model

Mixed Finite Element (MFE) method is one of many finite element methods which solves Darcys law and conservation equation to gain the scalar variable and flux vector simultaneously.

The reason for using MFE method to solve the forward flow model is that MFE can meet the following three excellent properties: local mass conservation, flux continuity, and the same convergence order for both the scalar variable and the flux. We firstly introduce some notations. Denote $(\cdot,\cdot)$ the $L^2(\Omega)$ inner product over a domain for scalar functions and $(L^2(\Omega))^2$ inner product for vector functions. Let us denote the following spaces by

$$\begin{aligned} V &= L^2(\Omega), \mathbf{W} = H(\mathrm{div};\Omega) = \left\{\mathbf{w} \in (L^2(\Omega))^2 : \nabla\cdot\mathbf{w} \in L^2(\Omega)\right\} \\ \mathbf{W}^0 &= \{\mathbf{w} \in H(\mathrm{div};\Omega), \mathbf{w}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega\}, \mathbf{W}_N^0 = \{\mathbf{w} \in H(\mathrm{div};\Omega), \mathbf{w}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega_N\}. \end{aligned}$$

Using these notations weak formulations for the flow model will be given in the subsections below followed by the MFE scheme for the flow model.

3.1 Weak formulation

The weak formulation is to seek $p \in V$ and $\mathbf{u} \in \mathbf{W}_N^0 + \mathbf{E}(u_B)$ such that

$$\left(K^{-1}\mathbf{u},\mathbf{w}\right) - (\nabla\cdot\mathbf{w}, p) = -\int_{\Gamma_D} p_B \mathbf{w}\cdot\mathbf{n}ds \text{ for all } \mathbf{w} \in \mathbf{W}_N^0 \tag{6}$$

$$(\nabla\cdot\mathbf{u}, v) = (q, v) \text{ for all } v \in V \tag{7}$$

where $\mathbf{E}(u_B)$ denotes a velocity extension such that $\mathbf{E}(u_B)\cdot\mathbf{n} = u_B$ on Γ_N.

3.2 MFE scheme

Generally, RT (Raviart-Thomas) finite element space is employed to approach the Darcy velocity. The r-th order RT space for the two-dimensional rectangular element is defined by

$$V_h(T) = P_r(T),\ \mathbf{W}_h(T) = (x_1P_r(T), x_2P_r(T)) \oplus (P_r(T))^2 \tag{8}$$

where $\oplus$ stands for the direct sum. $r = 0$ in our scheme is adopted in here. Accordingly, $\mathbf{W}_h(T)$ is given by

$$\mathbf{W}_h(T) = \{\mathbf{w} : \mathbf{w} = (a_T + b_Tx_1, c_T + b_Tx_2), a_T, b_T, c_T \in R\} \tag{9}$$

Restricted to an element T, $P_r(T)$ is the polynomial space of degree less than or equal to r. Consequently the MFE method for approximating the weak formulation is to search for $p_h(\cdot,\cdot) \in V_h$ and $\mathbf{u}_h \in \mathbf{W}_N^0 = \mathbf{W}^0 + \mathbf{E}_{u_B}$ such that

$$\left(K^{-1}\mathbf{u}_h,\mathbf{w}\right) - (\nabla\cdot\mathbf{w}, p_h) = -\int_{\Gamma_D} p_B \mathbf{w}\cdot\mathbf{n}ds \text{ for all } \mathbf{w} \in \mathbf{W}_h \tag{10}$$

$$(\nabla\cdot\mathbf{u}_h, v) = (q_h, v) \text{ for all } v \in V_h \tag{11}$$

MFE formulation causes a saddle point problem for the elliptic equations, the solution strategy is to use the Mixed-Hybrid algorithms for the pressure equation.

We consider a single rectangular element T. By RT_0 space, $\mathbf{u}_h$ in Equation (11) can be formulated as

$$\mathbf{u}_h = \sum_{E\in\partial T} \mathbf{u}_{T,E} w_{T,E} \tag{12}$$

where $w_{T,E}$ is a shape function of RT_0 and $\mathbf{u}_{T,E}$ is the total flux across an edge E of element T. Taking $w_{T,E}$ as the test function $\mathbf{w}$, the total flux is computed by

$$A_T U_T = p_T e - P_{TE} \tag{13}$$

and from Equation (13), we have

$$U_T = A_T^{-1}(p_T e - P_{TE}) \tag{14}$$

where $A_T = [(A_T)_{E,E' \in \partial T}]$, $(A_T)_{E,E'}) = \int_T w_{T,E} K^{-1} w_{T,E'} dT$, $U_T = [(U_T)_{E \in \partial T}]$, $e = [1]_{E \in \partial T}$, $P_{T,E} = [\overline{p}_{T,E}]$. Hence, the flows $u_{T,E}$ in Equation (13) passing through each edge is computed through a function of the element pressure average and edge average pressure. For simplicity, we reformulate the equation

$$u_{T,E} = a_{T,E} p_T - \sum_{E' \in \partial T} b_{T,E,E'} \overline{p}_{T,E'} \tag{15}$$

where $a_{T,E} = \sum_{E' \in \partial T} \left(A_T^{-1}\right)_{E,E'}$, $b_{T,E,E'} = \left(A_T^{-1}\right)_{E,E'}$.

For the second Equation of (11) and combining Equation (12), we obtain

$$B_T U_T = 0 \tag{16}$$

where $B_T = (B_T)_{T,E}$, $(B_T)_{T,E} = \int_T \nabla \cdot w_{T,E} dT, (w_T = 1)$.

Substituting Equation (14) into Equation (16) gives

$$B_T A_T^{-1} p_T e - B_T A_T^{-1} P_{T,E} = 0 \tag{17}$$

and therefore

$$p_T = (B_T A_T^{-1} e)^{-1} B_T A_T^{-1} P_{T,E}. \tag{18}$$

The continuity of the flux across the inter-element boundaries, $u_{T,E}$ leads to

$$u_{T,E} = \begin{cases} u_E^N, E = T \bigcap T' \\ -u_{T,E'}, E \in \Gamma_N \end{cases} \tag{19}$$

Finally, by using Equations (13), (18), and (19), an algebraic linear system of P_E is given as follows,

$$A_S P_E = J \tag{20}$$

where $A_S = D^S F - G$, $D = [a_{T,E}]_{N_T,N_E}, E \in \partial T$, $F = [(B_T A_T^{-1} e)^{-1} B_T A_T^{-1}]_{N_T}$, $G = [\sum_{E,E \in \partial T} b_{T,E,E'}]_{N_E,N_{E'}}$, $E \in \Gamma_D$, and N_T and N_E represent the number of elements and the number of edges not on Γ_D in the mesh, respectively. J is a vector of size N_E from boundary conditions.

4 The Inverse Problem

In the previous sections we saw how to find the measurements $\Upsilon = (\mathbf{u} \cdot \mathbf{n}, p)$ on the boundary $\partial\Omega$ for a given permeability K in a porous media Ω. The inverse problem is to find the permeability

K for given noisy measurements Υ^δ on the boundary $\partial\Omega$. The natural parameter space for the inverse problem is

$$Q = \{K \in L^\infty(\Omega) | 0 < K_{min} \leq K \leq K_{max} < \infty\}, \tag{21}$$

where $K_{min}, K_{max} > 0$ are constants. The inverse problem for the single phase Darcy flow problem in the deterministic setting for the infinite dimensional case can be formulated as a minimization problem. For example, a total variation (TV) regularization functional can be written as follows,

$$J_\alpha(K) = ||\Upsilon(K) - \Upsilon^\delta||_{L_2(\Omega)} + \alpha||\nabla K||_{L_1(\Omega)}, \tag{22}$$

where $\Upsilon(K)$ represents the solution on the boundary $\partial\Omega$ of the forward model and Υ^δ is the noisy experimental measurement and α is is a TV regularization parameter. Hence solving the inverse problem with this model would be equivalent of finding $\hat{K} = \arg\min_{K\in Q} J_\alpha(K)$.

Note that Υ^δ is a noisy measurement which implies that $\hat{K}$ won't be the real absolute permeability of the problem, it is rather the permeability with some kind of noise. This leads to the statistical inverse problem, which is finding the posterior density of the absolute permeability K. To simplify notation, we will denote $\underline{K}$ for the rest of this paper as the finite dimensional random variable of the absolute permeability and its concrete values at each pixel. Further, denote $\underline{\Upsilon}^\delta$ as the finite dimensional measurement. For example, the posterior density corresponding to the previous deterministic example (22) would be

$$P(\underline{K}|\underline{\Upsilon}^\delta) \propto \exp\left[-\frac{1}{2\sigma^2}||\underline{\Upsilon}(\underline{K}) - \underline{\Upsilon}^\delta||_{\ell_2}^2 - \alpha||\nabla\underline{K}||_{\ell_1}\right], \text{ for } \underline{K} \in \underline{Q}, \tag{23}$$

where σ^2 is the variance of the measurement noise and $\underline{Q} = \{\underline{K} \in \mathbb{R}^d | 0 < \underline{K}_{min} \leq \underline{K} \leq \underline{K}_{max} < \infty\}$. Note that $\exp[-\frac{1}{2\sigma^2}||\underline{\Upsilon}(K) - \underline{\Upsilon}^\delta||_{\ell_2}^2]$ is proportional to a multivariate normal distribution, which somehow implies that by using the ℓ_2 norm in the data term we assumed that $\underline{\Upsilon}^\delta$ has additive independent Gaussian noise. For the rest of this paper we will assume that $\underline{\Upsilon}^\delta$ has additive Gaussian noise. Hence, the data term for our problem is proportional to a multivariate normal density

$$P(\underline{\Upsilon}^\delta|\underline{K}) \propto \exp\left[-\frac{1}{2}(\underline{\Upsilon}(\underline{K}) - \underline{\Upsilon}^\delta)^T C^{-1}(\underline{\Upsilon}(\underline{K}) - \underline{\Upsilon}^\delta)\right], \tag{24}$$

where C is a positive definite covariance matrix of the measurement noise. Furthermore, we will use the more general regularization density

$$P(\underline{K}) \propto \begin{cases} \exp\left[-\alpha R(\underline{K})\right], & \text{if } \underline{K} \in \underline{Q}, \\ 0, & \text{else}, \end{cases} \tag{25}$$

where $R(\underline{K})$ is a regularization function. Concluding, the general posterior density becomes

$$P(\underline{K}|\underline{\Upsilon}^\delta) \propto P(\underline{\Upsilon}^\delta|\underline{K})P(\underline{K}). \tag{26}$$

By definition, the solution of the statistical inverse problem is the posterior density for the absolute permeability $\underline{K}$, but we are interested in finding the Bayes estimate $E(\underline{K}|\underline{\Upsilon}^\delta)$. Unfortunately, the posterior density (26) does not have a close form, which makes it impossible to find the Bayes estimate in a direct way. Therefore, we approximate the Bayes estimate via simulations.

5 The Markov Chain Monte Carlo Method

The idea of the MCMC method is generating a large random sample from the posterior density $P(\underline{\Upsilon}^\delta|\underline{K})$, with let's say Y samples, and then approximate the Bayes estimate by the sample mean, e.g.

$$E(\underline{K}|\underline{\Upsilon}^\delta) = \int_{\mathbb{R}^d} \underline{K} P(\underline{K}|\underline{\Upsilon}^\delta) d\underline{K} \approx \frac{1}{Y}\sum_{i=1}^{Y} \underline{K}_i, \tag{27}$$

where d is the number of dimensions of $\underline{K}$ (the number of pixels) and $\underline{K}_i$ represents the i^{th} random sample of the posterior density (26). Typical algorithms to generate such large random samples are the Gibbs sampler or the Metropolis Hastings algorithm. In this manuscript we will use a special type of an adaptive Metropolis algorithm. For detailed description of the Metropolis Hastings algorithm we recommend [4].

Let our Markov chain be defined on the continuous state space $(E, \mathscr{B}, \mathscr{M})$ where $\mathscr{B}$ is a Borel σ-algebra on E and $\mathscr{M}$ the normalized Lebesgue measure on E where $\underline{Q} \subseteq E \subseteq \mathbb{R}^d$ Additionally, assume $\xi(x; A)$ is a transition kernel for $\underline{K} \in E$, where $A \in \mathscr{B}$. The transition kernel $\xi(\underline{K}; A)$ denotes the probability of moving from a current state $\underline{K}$ to another state in the set A. We would like to find conditions of the transition kernel $\xi(\cdot;\cdot)$ such that it converges to an invariant distribution π. Here π represents the distribution of the posterior density $P(\underline{K}|\underline{\Upsilon}^\delta)$. After some analysis we can see that one transition kernel converging to the invariant distribution π is

$$\xi_{MH}(\underline{K}; A) := \int_A q(\underline{K}, \underline{K}^*)\alpha(\underline{K}, \underline{K}^*) d\underline{K}^* + \left[1 - \int_{\mathbb{R}^d} q(\underline{K}, \underline{K}^*)\alpha(\underline{K}, \underline{K}^*) d\underline{K}^*\right] \chi_A(\underline{K}), \tag{28}$$

where χ_A is the indicator function over the set A, $q(\underline{K}, \underline{K}^*)$ is a candidate-generating density, that is, a density which generates from a current random sample $\underline{K}$, a new candidate random sample $\underline{K}^*$. For example, $q(\cdot,\cdot)$ could be a multivariate normal density with mean $\underline{K}$. Note that the acceptance ratio

$$\alpha(x, y) = \begin{cases} \min\left[\frac{P(\underline{K}^*|\underline{\Upsilon}^\delta) q(\underline{K}^*, \underline{K})}{P(\underline{K}|\underline{\Upsilon}^\delta) q(\underline{K}, \underline{K}^*)}, 1\right], & \text{if } P(\underline{K}|\underline{\Upsilon}^\delta) q(\underline{K}, \underline{K}^*) > 0 \\ 1, & \text{otherwise,} \end{cases} \tag{29}$$

is the probability of accepting a new random sample $\underline{K}^*$. Note that in case that $q(\cdot,\cdot)$ is a symmetric density we have that $q(\underline{K}, \underline{K}^*) = q(\underline{K}^*, \underline{K})$ which simplifies (29). Concluding, the idea of the Metropolis Hastings algorithm is to generate new candidates of the absolute permeability using the proposal density $q(\cdot,\cdot)$ and then accept them as random sample of the posterior distribution with probability $\alpha(\underline{K}, \underline{K}^*)$.

It is known that the proper choice of the proposal density $q(\cdot,\cdot)$ for the Metropolis algorithm is vital to obtain a reasonable result by simulation in a suitable amount of time. This choice is generally very difficult since the target density is generally unknown [11, 8, 10]. One possible way of smoothing out this problem is by using an adaptive Metropolis algorithm which is iteratively updating the proposal density in an appropriate way. The downside of this kind of adaptive algorithm is that the chain usually becomes non-Markovian which would require to establish the correct ergodic properties. In other words, typically adaptive algorithms iteratively change the covariance matrix from the proposal density based on all previous samples of the chain, and hence $P(\underline{K}_n|\underline{K}_0, \underline{K}_1, ..., \underline{K}_{n-1}) \neq P(\underline{K}_n|\underline{K}_{n-1})$.

One adaptive algorithm, proposed in [10], changes the covariance matrix of the proposal distribution at atomic times in order to maintain the Markov property of the chain. An atomic time for a continuous state space $(E, \mathscr{B}, \mathscr{M})$ is a set $A \in \mathscr{B}$ with $\pi(A) > 0$ such that $\underline{K}_{n+1}, \underline{K}_{n+2}, ...$ is conditionally independent of $\underline{K}_0, \underline{K}_1, ..., \underline{K}_n$ given that $\underline{K}_n \in A$. Although this method seems very attractive at the first view it is practically very complicated to find proper atomic times for high dimensional problems [10]. Another approach is to adapt the proposal distribution a fixed amount of times and starting the burn-in time after the last adaptation. This method can not guarantee obtaining the optimal proposal distribution for the target distribution after the last adaptation. However, it usually increases the convergence speed considerable respect to the classical Metropolis-Hastings algorithm while still maintaining all good properties of it after the Burn-in time. Meaning the pilot adaptive Metropolis algorithm still generates a Markov chain after the pilot time. Another approach, introduced in [11], is to adapt the covariance matrix of a normal proposal distribution in every iteration after an initial time t_0 in such a way that the correct ergodic properties of the chain can be established even if the chain is itself non-Markovian. We will use a special kind of pilot adaptive Metropolis algorithm.

The idea of this algorithm is to train the proposal distribution by changing its covariance matrix in such a way that the acceptance ratio of the chain after the last adaption is close by the optimal acceptance ratio a_o of the chain. Note that there is no analytical framework for the choice of such a optimal acceptance ratio a_o when the target distribution is unknown. Therefore, the choice of a_o is usually based on the result from [8] that for a normal target and proposal distribution the optimal acceptance ratio is approximately .45 in the one dimensional case and .234 when the number of dimension converges to infinite.

Assume we wish to perform M adaptions, one every m iterations, where $1 < mM < B < B + Y$. Here B is the burn in time, meaning the amount of random simples which we consider to dependent from our starting guess and therefore not real random samples from the posterior density. Note that Y denotes the desired number of real random samples. Lets c_i denote a variable saving wether or not the i^{th} iteration of Algorithm 1 has been accepted,

$$c_i := \begin{cases} 1, & i\text{-th iteration has been accepted,} \\ 0, & \text{else.} \end{cases} \tag{30}$$

The estimator for the acceptance ratio for the j-th proposal distribution is $\bar{a}_j = \frac{1}{m} \sum_{i=(j-1)m+1}^{jm} c_i$. Let $0 < \epsilon \ll 1$, where 100ϵ would be the percentage of change per adaption in the covariance matrix C of the proposal distribution. In other words, the j-th adaption would modify the current covariance matrix C_{j-1} in the following way,

$$C_j = \Xi_{PAM}(C_{j-1}) := \begin{cases} (1+\epsilon)C_{j-1}, & \text{if } \bar{a}_j > a_o, \\ C_{j-1}, & \text{if } \bar{a}_j = a_o, \\ (1-\epsilon)C_{j-1}, & \text{if } \bar{a}_j < a_o. \end{cases} \tag{31}$$

Informally speaking, the algorithm would modify the covariance matrix in the pilot time mM in such a way that it comes closer to one which has an optimal acceptance ratio, and then starts the standard Metropolis Hastings algorithm with the latest state and proposal distribution of the pilot time. In Algorithm 1 we recapitulate the pilot adaptive Metropolis algorithm with an arbitrary starting state $\underline{K}_0 \in \underline{Q}$ and a starting guess for the positive definite covariance matrix C_0.

When applying Algorithm 1 it is important to know that the chain only satisfies the Markov property after the last adaption at time mM. Note that the chain will still move towards the high probability areas of the target distribution during the pilot time, which usually results in a considerable shorter Burn-in time $B > mM$ comparing to the standard Metropolis Hastings algorithm.

Algorithm 1 A Pilot Adaptive Metropolis Algorithm.

$j = 1$;
for i = 1 to B+Y **do**
  if $i \equiv 0 \bmod m$ and $i \leq mM$ **then**
    $C_j = \Xi_{PAM}(C_{j-1})$;
    j++;
  end if
  Generate $\underline{K}^*$ from $q_{C_j}(\underline{K}_{i-1}, \cdot)$ and u from $\mathcal{U}(0,1)$;
  if $u \leq \alpha(\underline{K}_{i-1}, \underline{K}^*)$ **then**
    $\underline{K}_i = \underline{K}^*$;
  else
    $\underline{K}_i = \underline{K}_{i-1}$;
  end if
end for
Return $\left\{\underline{K}_1, \underline{K}_2, ..., \underline{K}_{B+Y}\right\}$

We discussed the Metropolis-Hastings algorithm and we saw that it samples properly from the posterior distribution only after a burn-in time B. Therefore, we are interested in knowing how long this burn-in phase should be. Unfortunately, there is no theory giving a good estimate of the burn-in time prior to run the Metropolis-Hastings algorithm. That leaves us to first run the Metropolis-Hastings algorithm and then check wether or not it converged to its stationary distribution (Invariant distribution). There are several methods to check whether or not the chain converged, however for all methods have a positive probability to produce give a wrong answer. Therefore, it is recommended to use more that one diagnostic method. The most common diagnostic methods in the literature are found in the work of Gelman and Rubin [9, 3], Geweke [6] and Raftery and Lewis [15].

6 Regularizing Prior Density Selection

Before presenting simulations we need to explain what kind of regularizing function $R(\cdot)$ in (25) we are using. A common choice for regularization in a deterministic setting is a penalty $R(\underline{K})$ using a ℓ_p norm with $1 \leq p \leq 2$. For statistical regularization, there is no need to assume that $R(\underline{K})$ be neither a norm nor a convex function, we only need it to be a continuous function. Therefore we can write the following more general regularization:

$$R_1(\underline{K}) := \sum_{i=1}^{M} |\underline{K}_i - \underline{K}_i^{pr}|^s,$$

where $0 < s \leq 2$, $\underline{K}^{pr}$ is a prior estimate of $\underline{K}$. In case that $0 < s < 1$ this only defines a metric but not a norm. The regularization function R_1 would make $\underline{K}$ smooth and is specially efficient for reconstructions where the area different from the background is small. Another common choice would be the total variation prior, which is defined as:

$$R_2^*(K) = ||\nabla K||_{L_1(\Omega)} = \int_\Omega |\nabla K| dx.$$

Let $R_2(\underline{K})$ represent the discrete analogue of $R_2^*(K)$, for an exact formulation see [14]. This prior makes the image of the $\underline{K}$ smooth.

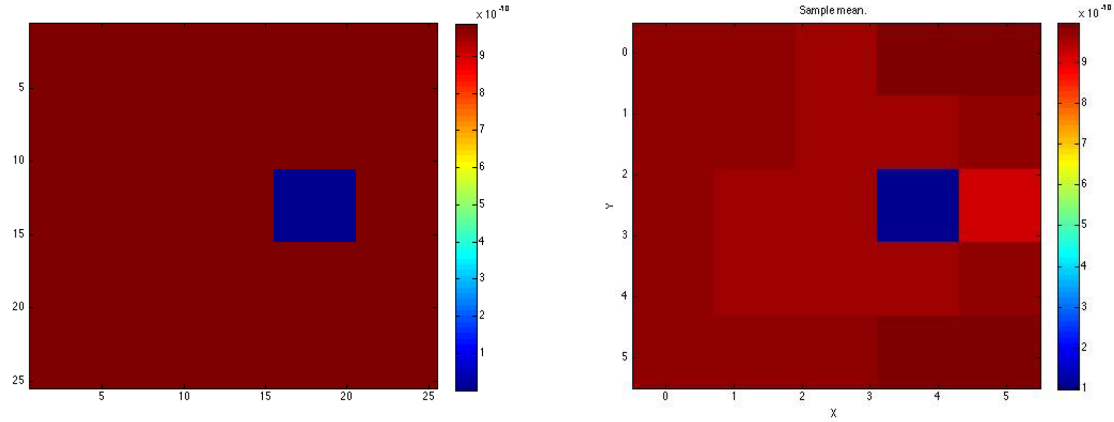

Figure 1. The Darcy flow is from left to right. The image on the left hand side represents the true absolute permeability on a 25×25 mesh. The image on the right hand side is a reconstruction using algorithm 1 with a R_1 (ℓ_1) regularization on a 5×5 mesh. Note that $\frac{||\underline{K}_{true}-\underline{K}_{rec}||_{L_1}}{||\underline{K}_{true}||_{L_1}} \approx 0.023$ where $\underline{K}_{true}$ is the true permeability and $\underline{K}_{rec}$ is the reconstructed permeability.

7 Preliminary Simulation and Conclusion

Here we simulate measurements of a single phase Darcy flow (going from left to right see Figure 1) on a 25×25 mesh for the velocity, pressure, and absolute permeability for the forward problem and then we use algorithm 1, with the posterior density introduced in (26) and a R_1 regularization, to reconstruct $\underline{K}$ on a 5×5 mesh. We note that this is a typical setup to avoid inverse crimes mainly the data using the forward problem is generated on a different mesh than the mesh used for the inverse problem. In algorithm 1 we use the pilot time $mM = 70,000$ with one adaption every $m = 100$ iteration, we choose the Burn-in time to be $B = 140,000$ and the total number of real random sample $Y = 80,000$. The starting absolute permeability $\underline{K}_0$ has been chosen to be constant. The starting covariance matrix C_0 was proportional to a identity matrix plus an additional covariance of .1 for every neighbor elements of the mesh. The result of the image reconstruction can be found in Figure 1. For this particular simulation which is low dimensional with $d = 25$, the computational time is in the order of minutes, however in realistic scenario for a permeability with much higher dimension i.e. large d, the inverse problem gets computationally very intensive. This is one of the drawbacks of the statistical approach using MCMC based algorithms. We have also used deterministic methods in MATLAB such as Levenberg-Marquardt or Gauss-Newton for the reconstruction of $\underline{K}$ and found that the statistical approach provides a much better image. However, further investigation is required for a thorough comparison. In the future, we will be performing more simulations for different shapes for $\underline{K}$ as well as implement the TV penalty term and explore the mixture of TV and sparsity regularization.

We conclude that statistical inversion approach using Markov Chain Monte Carlo method using an adaptive Metropolis-Hastings algorithm can be used to regularize the ill-posed inverse problem of permeability reconstruction in porous media.

References

[1] John Baird and Clint Dawson. The representer method for data assimilation in single-phase darcy flow in porous media. *Computational Geosciences*, 9(4):247–271, 2005.

[2] Johnathan M Bardsley. Mcmc-based image reconstruction with uncertainty quantification. *SIAM Journal on Scientific Computing*, 34(3):A1316–A1332, 2012.

[3] Stephen P Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics*, 7(4):434–455, 1998.

[4] Siddhartha Chib and Edward Greenberg. Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327–335, 1995.

[5] Clint Dawson, Shuyu Sun, and Mary F. Wheeler. Compatible algorithms for coupled flow and transport. *Computer Methods in Applied Mechanics and Engineering*, 193(2326):2565 – 2580, 2004.

[6] Salaheddine El Adlouni, Anne-Catherine Favre, and Bernard Bobée. Comparison of methodologies to assess the convergence of markov chain monte carlo methods. *Computational Statistics & Data Analysis*, 50(10):2685–2701, 2006.

[7] Xiaolin Fan, Shuyu Sun, Wei Wei, and Jisheng Kou. Numerical simulation of pollutant transport in fractured vuggy porous karstic aquifers. *Journal of Applied Mathematics*, 2011, 2011.

[8] A Gelman, G Roberts, and W Gilks. Efficient metropolis jumping rules. *Bayesian statistics*, 5:599–608, 1996.

[9] Andrew Gelman and Donald B Rubin. Inference from iterative simulation using multiple sequences. *Statistical science*, pages 457–472, 1992.

[10] Walter R Gilks, Gareth O Roberts, and Sujit K Sahu. Adaptive markov chain monte carlo through regeneration. *Journal of the American statistical association*, 93(443):1045–1054, 1998.

[11] Heikki Haario, Eero Saksman, and Johanna Tamminen. An adaptive metropolis algorithm. *Bernoulli*, pages 223–242, 2001.

[12] Marco A Iglesias and Clint Dawson. The representer method for state and parameter estimation in single-phase darcy flow. *Computer Methods in Applied Mechanics and Engineering*, 196(45):4577–4596, 2007.

[13] Bangti Jin, Taufiquar Khan, and Peter Maass. A reconstruction algorithm for electrical impedance tomography based on sparsity regularization. *International Journal for Numerical Methods in Engineering*, 89(3):337–353, 2012.

[14] Jari P Kaipio, Ville Kolehmainen, Erkki Somersalo, and Marko Vauhkonen. Statistical inversion and monte carlo sampling methods in electrical impedance tomography. *Inverse problems*, 16(5):1487, 2000.

[15] Adrian E Raftery and Steven M Lewis. Implementing mcmc. In *Markov chain Monte Carlo in practice*, pages 115–130. Springer, 1996.

[16] Shuyu Sun. Discontinuous galerkin methods for reactive transport in porous media. 2003.

[17] Shuyu Sun and Mary F Wheeler. Discontinuous galerkin methods for coupled flow and reactive transport problems. *Applied Numerical Mathematics*, 52(2):273–298, 2005.

[18] Shuyu Sun and Mary F Wheeler. $L^2(H^1)$ norm a posteriorierror estimation for discontinuous galerkin approximations of reactive transport problems. *Journal of Scientific Computing*, 22(1-3):501–530, 2005.

[19] Shuyu Sun and Mary F Wheeler. Anisotropic and dynamic mesh adaptation for discontinuous galerkin methods applied to reactive transport. *Computer methods in applied mechanics and engineering*, 195(25):3382–3405, 2006.

[20] Shuyu Sun and Mary F. Wheeler. Symmetric and nonsymmetric discontinuous galerkin methods for reactive transport in porous media. *SIAM Journal on Numerical Analysis*, 43(1):pp. 195–219, 2006.