VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

School of Engineering and Computer Science
Te Kura Mātai Pūkaha, Pūrorohiko

PO Box 600
Wellington
New Zealand

Tel: +64 4 463 5341
Fax: +64 4 463 5045
Internet: office@ecs.vuw.ac.nz

Bayesian estimation of bound fluid fraction from NMR relaxation Final Report

David Dobbie

Supervisors: Paul Teal, Robin Dykstra

Submitted in partial fulfilment of the requirements for
Bachelor of Engineering with Honours in Electronic and
Computer Systems Engineering.

Abstract

Effective measurements of NMR relaxation have typically been constrained to high SNR environments due to the adverse effect that noise can have. Inversion of the measured data to its relaxation density function is an ill-posed problem. The current literature has worked on increasing the accuracy of inversion given only measured data and the nature of the noise. These techniques have been recreated and bench marked in this report. This project seeks to use Bayesian techniques to integrate prior data based on high quality measurements in order to make more accurate inversions of measured data. This is used to compute the bound fluid fraction of porous media to assess the difficulty of fluid extraction.

Contents

Chapter 1

Introduction

Relaxation times gathered from nuclear magnetic resonance (NMR) experiments provide insight into the physical nature of a sample [1]. However, most NMR experiments are restricted to a limited and controlled environment - like a laboratory [2]. To be practicable in use cases that have hard constraints such as high noise or low signal power, more advanced signal processing is required. This report details a new technique which extends of the use of NMR in noisy environments by utilising prior high quality measurement data.

1.1 Motivation

The focus of this project is to create a more noise resistant estimator of the bound fluid fraction (BFF) of porous media via NMR relaxation. This physical property is crucial for evaluating the viability of an oil well [3]. Fluid that is bound to a porous media, sandstone for example, is considered uneconomic to obtain in a mineral extraction process. The noisy environment of the oil well complicates estimation adversely [3]. This necessitates the use and design of robust estimation techniques [4][5], such as the Bayesian technique proposed by this project.

Beyond the estimation of BFF, the improvements on NMR signal processing made in this project are potentially transferable to other applications where power is limited. These can include identifying the nature of fluids in food products so that we may minimise wastage [6]. Therefore, a robust analysis of the proposed estimator will form an indication of further viability.

1.2 Goals

The goals of this project were to:

1. implement and evaluate the competing techniques for estimating the BFF of a sample via T2 relaxation measurements;

2. design, implement, and evaluate a new technique using Bayesian statistics to estimate the BFF of a sample via T2 relaxation.

1.3 Report Outline

This report discusses in detail the current field of techniques and the method proposed by this project. The structure of it is as follows:

- Chapter 2 - Background - establishes the physical basis of NMR relaxation for bound fluid fraction and explores the current literature for estimation of the desnity of relaxation times.
- Chapter 3 - Design - explores the design aspects of the proposed technique such as constraints and pathways chosen.
- Chapter 4 - Implementation - explores the development of the original techniques and the proposed technique, the validation of these techniques, and the test architecture created for comparison.
- Chapter 5 - Evaluation - directly compares the techniques' performance, with attention paid towards performance for a typical operating point in the field.
- Chapter 6 - Conclusions - summarises this report and discusses possible future pathways beyond this project.

Chapter 2

Background

Processing NMR relaxation in low SNR environments is crucial for dynamic environments such as well logging [3] [7]. The previously published work forms the benchmark for comparison with this project's proposed technique. This background survey details the current published technique, the evaluation criteria of the techniques, and the Bayesian theory that informs the design in Chapter 3.

2.1 NMR Relaxation

Nuclear magnetic resonance involves determining the nature of the nuclei of a sample by using its magnetic behaviour [8]. To provide an effective basis for the signal processing techniques to operate on, we start by examining the physics of the magnetism of the nuclei.

2.1.1 Underlying Physics

Magnetism of the Nuclei

The nucleus has two classical aspects that make it a microscopic magnet [8]:

1. a nucleus has an electrical charge from the protons that make it up, and
2. it may have a non-zero spin.

The spin creates a magnetic dipole moment, $\vec{\mu}$. This phenomenon is expressed as the relation of angular momentum $\vec{J}$ and the gyro-magnetic ratio inherent to the nuclei, γ, such that

$$\vec{\mu} = \gamma \vec{J} \tag{2.1}$$

A macroscopic magnetic field formed by many nuclei, $\vec{M}$, is the summation over the total number of spins in the system, N_s [8].

$$\vec{M} = \sum_{n=1}^{N_s} \vec{\mu_n} \tag{2.2}$$

Introducing a static magnetic field B_0 induces rotation of the magnetic moments of the nuclei. This rotation about the axis of B_0 is called nuclear precession. The angular frequency of this rotation is known as the Larmor frequency [8]. This is given by:

$$\omega_0 = \gamma B_0 \tag{2.3}$$

Holding the magnetic field B_0 constant means that all the same nuclei exhibit nuclear precession at the same frequency. Their magnetic dipoles precess at the same frequency.

In order to set all of the rotations of magnetic dipoles into the same phase we require another magnetic field, B_1. This field makes the system phase coherent [8]. This is so that all of the magnetic dipoles of the same frequency constructively interfere, providing a strong enough signal to measure. B_1 is in the form of a short RF (radio frequency) pulse that lasts from microseconds to the milliseconds [8].

With phase coherence achieved there is a measurable macroscopic magnetic field originating from the nuclei. The change of this nuclei magnetic field is described by the Bloch Equation [8]:

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B} - \frac{M_x\vec{i} + M_y\vec{j}}{T_2} - \frac{(M_z - M_z^0)\vec{k}}{T_1} \tag{2.4}$$

where $\vec{M} = (M_x, M_y, M_z)$.

Relaxation

At the end of an RF pulse, the magnetic field B_1 is set to 0 [8]. In the rotating frame of reference this eliminates the first term of the Bloch equation. This results in the magnetism of the system of nuclei to be described by two differential equations:

$$\begin{cases} \frac{dM_{z'}}{dt} = -\frac{M_{z'} - M_z^0}{T_1} \\ \frac{dM_{x'y'}}{dt} = -\frac{M_{x'y'}}{T_2} \end{cases} \tag{2.5}$$

The time evolution of the system is governed by these equations. Immediately after the RF pulse ($t = 0_+$) for the transverse magnetic field, $M_{x'y'}$, is expressed as:

$$M_{x'y'}(t) = M_{x'y'}(0_+)e^{-t/T_2} \tag{2.6}$$

This exponential decay is referred to as the T_2 relaxation [8]. The decay constant T_2 is used to give insight to the physical properties of the nuclei in the sample. This exponential model is only suitable for weak spin-spin systems such as fluids. Therefore, the analysis of the T_2 relaxation is limited to the behaviour of the fluids in a sample.

Detecting Relaxation

The resultant magnetic field can be detected by using Faraday's Law and the principle of reciprocity [8]:

$$V(t) = -\frac{\partial\Phi(t)}{\partial t} = -\frac{\partial}{\partial t}\int_{\text{object}} \vec{B_r}(\boldsymbol{r}) \cdot \vec{M}(\boldsymbol{r} \cdot t)d\boldsymbol{r} \tag{2.7}$$

$\vec{B_r}(\boldsymbol{r})$ is the magnetic field produced by the coil by a hypothetical unit direct current. After demodulation of the voltage [8], we get a measurement signal we can process and analyse:

$$m(t) = \int_{\text{object}} M_{x,y}(\boldsymbol{r}, 0)e^{i\gamma\Delta B(\boldsymbol{r})t}d\boldsymbol{r} \tag{2.8}$$

The term ΔB varies in space due to the inhomogeneity of the magnetic field. This is a form of error introduced into the measurement process, which is more pronounced in field measurement tools rather than in laboratory tools. Figure 2.1 illustrates an example field measurement tool and its grossly inhomogeneous magnetic field.

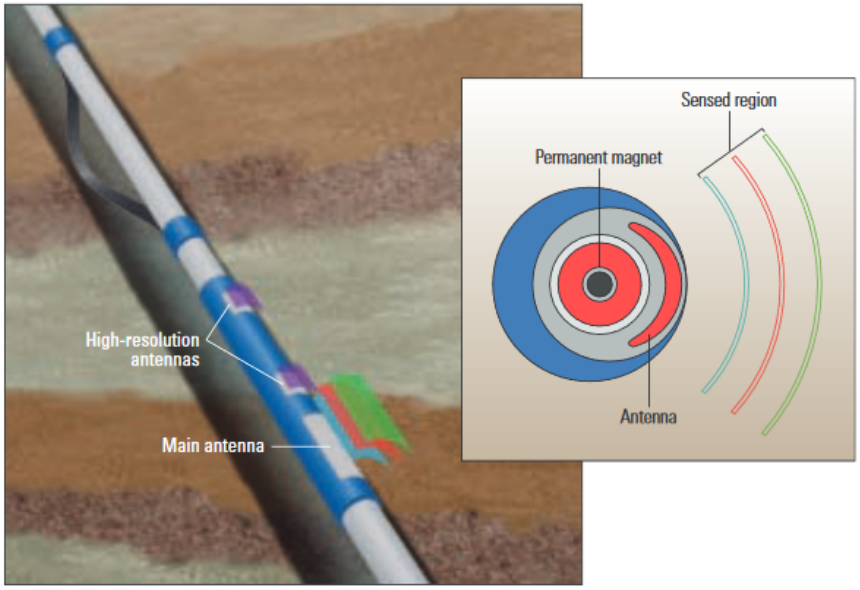
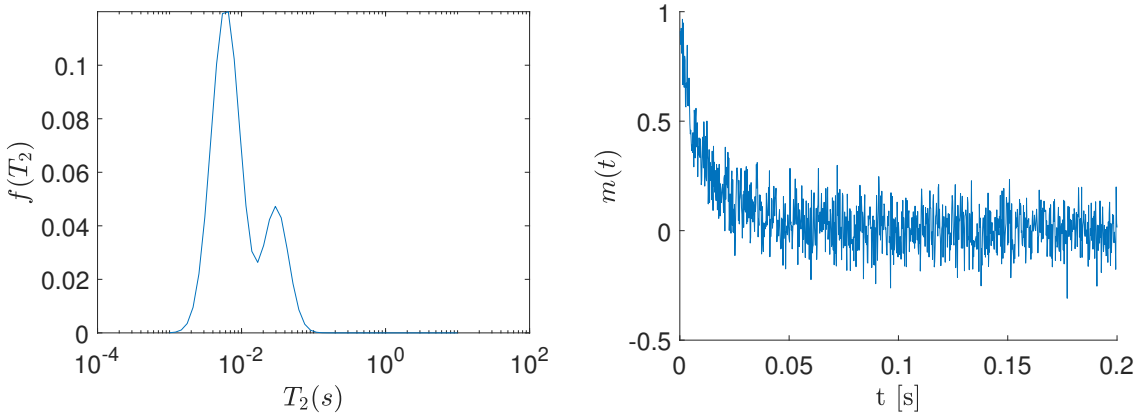


Figure 2.1: An example of a field measurement tool used to detect relaxation for the rock in a bore hole [9]

.

2.1.2 The T_2 Density Function

The density function of T2 relaxation times is denoted by $f(T_2)$. An example T_2 distribution is shown in Figure 2.2. The T2 density function forms the basis of analysis of fluids in a sample. There are several properties that relate T2 relaxation to fluids in porous media. The most significant two we shall analyse are: *porosity*, and *bound fluid fraction* [10] [3].



(a) The density function $f(T_2)$

(b) The modelled measured data $m(t)$

Figure 2.2: An example modelled T_2 density function and the measured data generated from it. It has unity porosity and a noise standard deviation of $\sigma_\epsilon = 0.2$

Porosity

The fraction of the rock volume that can be occupied by fluid is defined as the porosity (ϕ) [11]. After calibration with a water sample ($\phi = 1$), the integral over the entire density

function is equal to the porosity.

$$\phi = \int_0^\infty f(T_2) dT_2 \tag{2.9}$$

This is approximated by taking the sum over the discretised density function:

$$\phi \approx \sum_i f(T_{2i}) \tag{2.10}$$

Bound Fluid Fraction

Bound fluids are fluids that require significant capillary pressure to remove from porous media [3]. This pressure is caused by the fluid being located in small pores of the media. The volume of bound fluids – the bound fluid volume (BFV) [1] – is associated with relaxation times below a cut off value, T_c [3]. This volume is obtained with:

$$BFV = \int_0^{T_c} f(T_2) dT_2 \tag{2.11}$$

Free fluids conversely are fluids that are able to be extracted since they experience less capillary pressure restraining them [10][12]. Therefore, the free fluid volume – also known as movable fluids (BVM) – can be expressed as the remaining fluid that is not bound, having its T2 relaxation above the cut-off. Since the porosity is defined to be all of the fluid volume in the sample, the bound fluid fraction (BFF) can be expressed in terms of the porosity.

The bound fluid fraction is expressed, as:

$$BFF = \frac{BFV}{BFV + BVM} = \frac{BFV}{\phi} = \frac{\int_0^{T_c} f(T_2) dT_2}{\int_0^\infty f(T_2) dT_2} \tag{2.12}$$

2.2 Inversion of Measured Data

In order to compute the T2 relaxation times from the measured data we require a computational model that corroborates with the physical model. This model reveals that the measured data is ill-posed for an inversion problem.

2.2.1 Model of T2 Relaxation Measurements

The measured T2 relaxation of a sample can be modelled with the following integral

$$m(t) = \int e^{\frac{-t}{T_2}} f(T_2) dT_2 + \epsilon \tag{2.13}$$

As we are unable to numerically compute continuous functions, we discretise this into

$$\hat{m}(t) = \sum_{k=1}^{N_y} f_k e^{\frac{-t}{T_{2_k}}} + \epsilon \tag{2.14}$$

This model corresponds with the detected exponential decay of the transverse magnetic field in (2.6). The measurement data m is made up of N_2 discrete measurements. The time constants of these exponential functions – T_{2_k} – make up the T2 relaxation times. In the model itself N_y is the total number of relaxation times modelled, f_k is the contribution of

[1] also known as the bulk fluid irreducible (BVI)

each exponential, and ϵ is the additive white Gaussian noise introduced by the measurement process.

Since we know the measurements are made of exponentially decaying functions, we can construct an exponential kernel matrix, K, that maps from the T2 domain to the time domain, i.e.

$$(K)_{ij} = e^{-t_i/T_{2,j}} \tag{2.15}$$

Hence, the T2 density function is modelled as a vector f. The overall model of the measurement process is described as:

$$m = Kf + \epsilon, \text{ where } m \in \mathbb{R}^{N_2},\ f \in \mathbb{R}^{N_y},\ K \in \mathbb{R}^{N_2 \times N_y},\ \epsilon \in \mathbb{R}^{N_2} \tag{2.16}$$

2.2.2 Ill-posed Nature

The goal of this project is to find the bound fluid fraction. This has traditionally required estimating f and then computing from there with an integral via (2.12). Transforming a *noiseless* single exponential to an easily quantifiable time constant can be achieved with the Inverse Laplace Transform (ILT) [13].

$$f_k \delta\left(\frac{1}{T_2} - \frac{1}{T_{2_k}}\right) = \mathcal{L}^{-1}\left\{f_k e^{\frac{-t}{T_{2_k}}}\right\} \tag{2.17}$$

However, this non-robust version of the ILT is not suited for the noisy model. There can be very different results for the same measured sample due to the random effect of noise. Also, the kernel matrix K is very poorly conditioned for computation [14]. The measured signal cannot be frequency filtered since that process does not discriminate between the noise or the true signal. This necessitates the implementation of a robust numerical approximation of the inversion of measurement data.

2.3 Existing Techniques

The typical design limitation of the existing techniques is their limit on prior information for the sake of generality. The main thrust of these previous designs to add more ways of evaluating error without relaxing the constraint on useful prior information. The preferred form of deriving the density function is also via an iterative computation rather than a direct analytic approach.

2.3.1 Technique Constraints

The existing techniques have constraints that limit their accuracy in estimation in exchange for generality. These include:

- knowing only the measurement, m, and the noise power of it, σ,
- that the density function f is non-negative,
- that the measurement is in the form of a decaying exponential, and
- the sample period of the measurements is known.

These constraints imply that more detailed prior information may allow for a greater reduction in error in comparison to the existing techniques. This potential performance increase is fundamental for the design of the Bayesian estimator in Chapter 3.

2.3.2 The Inverse Laplace Transform (ILT) Technique

Numerical inversion is based on the goal of minimising the difference between the true unknown distribution and the estimated distribution. The most popular inversion technique of T2 relaxation is that described by Venkataramanan et al.. for one dimensional and two dimensional distributions [15] [2]. This technique requires knowing the standard deviation of the noise present, σ_ϵ.

Inversion With Optimisation

The Bulter-Reeds-Dawson (BRD) method [13] provides the established technique for robust numerical computation of the ILT. This is an optimisation framework expressed as:

$$\Phi(f) = \min_{f \geq 0} ||Kf_\alpha - m||^2 + \alpha ||f_\alpha||^2 \tag{2.18}$$

In this framework we can see that the main cost being minimised is the difference between the hypothesised noiseless time domain data Kf_α and the actual measured data m. However, constricting the problem simply to just this allows for overfitting to the noise itself. To counteract this, the second term $\alpha||f_\alpha||^2$ is added to penalise the size of the hypothesised density function. This 'smooths' the hypothesised f to prevent it fitting to the noise [16]. This process is known as *regularisation*. The BRD method [13] sets the optimal regularisation parameter as

$$\alpha_{\text{opt}} = \frac{N_y \sqrt{\sigma_\epsilon}}{||c||}, \tag{2.19}$$

where the vector c satisfies the following simultaneous equations [13]:

$$f_{\text{est}} = \max(0, Kc) \tag{2.20}$$

$$(Kf - g) + \alpha c = 0 \tag{2.21}$$

This method involves updating c with α_{opt} and then computing α_{opt} for the new c (Figure 2.3). This iterative computation converges towards the density function we are trying to predict after approximately 20 iterations. The problem is convex [13], so the starting α and c can be set to any value as it will automatically converge to the optimal solution.

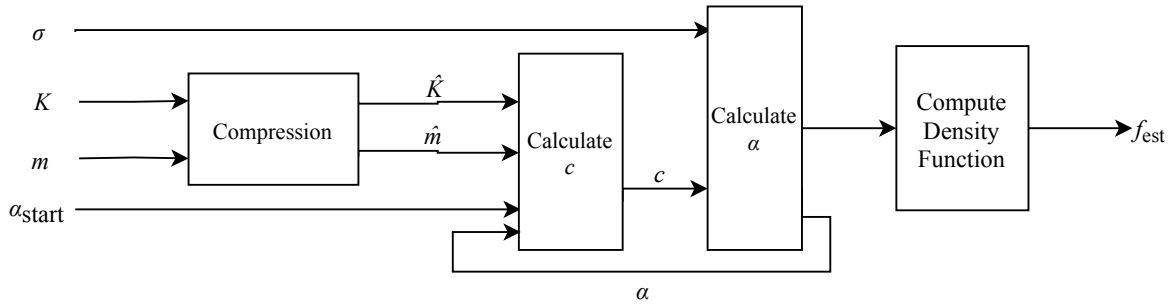


Figure 2.3: Computational work flow of the optimisation framework in the inverse Laplace transform approximation

[2] The ILT technique explored here is technically solving Fredholm integrals of the first kind.

Calculating c

Computing the c vector can be computationally dangerous due to potential error introduced by the inversion of a possibly ill behaved matrix [13]. The BRD method circumvents this problem by ensuring inversion is done by a positive semidefinite matrix:

$$H = \text{step}(KK^T) \text{ , where } \quad \text{step}(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{2.22}$$

that feeds in the regularisation coefficient [13] α to give

$$c = (KHK^T + \alpha I)^{-1} m. \tag{2.23}$$

Measurement Data Compression

To reduce computation time, truncated singular value decomposition of the kernel is used [15]. Starting with the singular value decomposition $K = USV^T$:

- the first n columns $\hat{U}$ and $\hat{V}$ are kept from U and V, and
- the first n columns and rows of $\hat{S}$ are kept from S [17].

$$\hat{K} = \hat{S}\hat{V}^T \tag{2.24}$$

$$\hat{m} = m^T \hat{U} \tag{2.25}$$

These are then used to make a new kernel and measurement vector. Here, the dimensionality is reduced from N_2 (thousands) dimensions to n dimensions. The optimisation problem remains unchanged as $||K||^2 = ||USV^T||^2 \approx ||SV^T||^2$; we are typically minimising the same error. This holds because the singular values of K decay very rapidly as it is an exponential kernel [5].

2.3.3 Direct Tapered Area Estimation Technique

Tapered areas are weighted areas under a T2 distribution [18]. This is a property of the density function that corresponds to bound fluid volume and porosity. Therefore, it is a viable candidate for estimating the bound fluid fraction of a sample directly from measurement data [1].

Tapered areas are computed directly from a density function with the use of a tapered step function. The specific transform used by Gruber et al. is the Exponential Haar Transform (EHT) [1]. This is defined in the T_2 domain as:

$$K_a(T_2, T_c) = \frac{C}{\gamma} \tanh(\alpha\gamma) \tag{2.26}$$

In addition, the EHT is defined in the time domain as:

$$k_a(t, T_c) = C(-1)^n e^{\beta t}, \quad 2n\alpha < t < 2(n+1)\alpha, \quad n \in \mathbb{Z} \tag{2.27}$$

The actual tapered area (B) and the estimate tapered area ($\hat{B}$) are computed with an inner product [3] of the EHT kernel and the data for their respective domains:

$$B = K_a(T_2, T_c) f(T_2), \quad \hat{B} = k_a(t, T_c) m(t) \tag{2.28}$$

[3] This is equivalent to an integral transform for the continuous case.

The transform in the T_2 domain is normalised so that where $T_2 = T_c$, $K = 0.5$ [1]. To meet this property, the constants are set by Gruber et al. to be:

$$C = \frac{0.7213}{T_c}, \quad \alpha = (1.572)T_c, \quad \beta = \frac{0.4087}{T_c}, \quad \gamma = \frac{1}{T_2} + \beta \tag{2.29}$$

2.3.4 The ILT+ Technique

More recent techniques have refined the optimisation approach with additional measures of evaluating estimation error. These come in the form of more prior information used to compensate for the high noise environment of the measured data [19]. This involves calculating aspects that are intrinsic to the T2 density function. These are known as linear functionals. The two main functionals used for this improved optimisation framework are *moments* and *tapered areas* (Figure 2.4) [19].

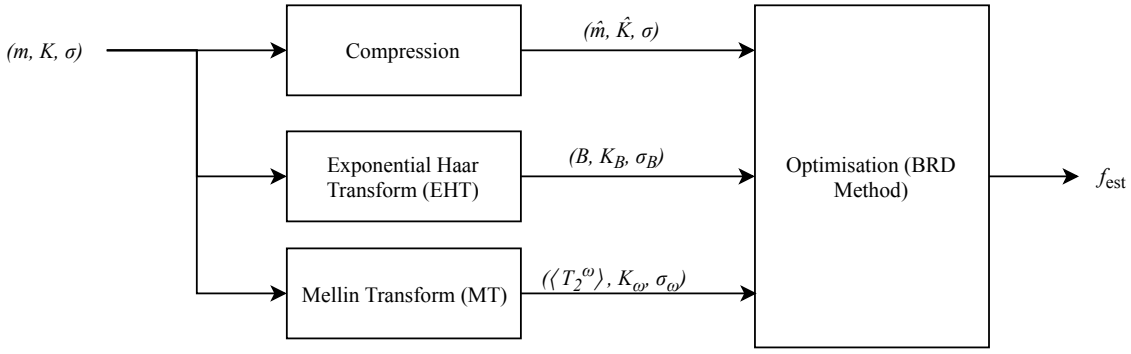


Figure 2.4: Computational work flow of the optimisation framework for the ILT+ method.

Estimation of Moments

A moment is a description of the 'weight' of the density function on the T2 axis [20]. An example of a moment is the first moment, the mean. The ω^{th} moment is defined as:

$$\langle T_2^{\omega} \rangle \equiv \frac{\int_0^{\infty} T_2^{\omega} f_{T_2}(T_2) dT_2}{\int_0^{\infty} f_{T_2}(T_2) dT_2} = \frac{\int_0^{\infty} T_2^{\omega} f_{T_2}(T_2) dT_2}{\phi} \tag{2.30}$$

The ILT+ requires an estimation of the moment directly from the measured data without knowledge of the density function. This estimation is implemented using the Mellin transform (MT) [20]:

$$\langle \hat{T_2^{\omega}} \rangle = \begin{cases} \frac{-1}{\Gamma(\mu)\phi} \int_0^{\infty} t^{\omega-1}[m(t) - \phi]dt & -1 < \omega < 0, \\ 1 & \omega = 0, \\ \frac{1}{\Gamma(\mu)\phi} \int_0^{\infty} t^{\omega-1} m(t) dt & \omega > 0, \end{cases} \tag{2.31}$$

The computed moment is a numerical value describing the density function. This has a propagated uncertainty described in [20] that can be used to evaluate how much we should depend on it for the optimisation framework.

Estimation of Tapered Areas

This linear functional is discussed in Section 2.3.3. The implementation is identical with the only difference being that the result is fed into the ILT+ framework.

Constrained Optimisation

Directly estimated linear functionals are introduced as additional methods of evaluating the estimate density functions' error [19]. These estimated functionals are also more accurate than simply using the ILT method [1]. Therefore, the additional prior information yields a more accurate estimation of the density function in a low SNR environment. This forms the basis for the ILT+ method [19]. The optimisation framework that aims to minimise the cost $Q(f_\alpha)$ takes the form of:

$$Q(f_\alpha) = \min_{f \geq 0} ||W(Lf_\alpha - g)||^2 + \alpha ||f_\alpha||^2 \tag{2.32}$$

This framework is adapted from the ILT method (eq. 2.18) by appending the estimations and uncertainties of the tapered areas and moments onto the framework [19] such that:

$$g = \begin{bmatrix} \hat{m} \\ \langle T_2^{\omega_1} \rangle \\ \vdots \\ \langle T_2^{\omega_{N_m}} \rangle \\ B_1 \\ \vdots \\ B_{N_a} \end{bmatrix}, \quad L = \begin{bmatrix} \hat{K} \\ \frac{1}{\phi} T_2^{\omega_1} \\ \vdots \\ \frac{1}{\phi} T_2^{\omega_{N_m}} \\ K_a(T_2, T_{c_1}) \\ \vdots \\ K_a(T_2, T_{c_3}) \end{bmatrix}, \quad W = \begin{bmatrix} \frac{1}{\sigma_\epsilon} \\ \frac{1}{\sigma_{\omega_1}} \\ \vdots \\ \frac{1}{\sigma_{\omega_{N_m}}} \\ \frac{1}{\sigma_B} \\ \vdots \\ \frac{1}{\sigma_{B_{N_a}}} \end{bmatrix} I \tag{2.33}$$

The g vector contains the compressed measurement, N_m estimated moments and N_a estimated tapered areas. The L matrix maps from the T2 domain to the time domain: the estimated density function (f), estimated moments of f, and estimated tapered areas of f. The W matrix is a diagonal matrix that changes the weight of each respective value of g and L using division by their respective uncertainty. Therefore, the larger weight value we have, the more certain we are that the estimation is correct. The m vector is compressed via the methodology detailed in Section 2.3.2. The SVD compression is used only for the measurement data and its respective kernel.

2.3.5 Existing Techniques' Summary

The published techniques all constrain their prior information to only the constraints detailed in Section 2.3.1. Any extra degrees of evaluating error are confined to tapered area and moment estimations that still constrain themselves to these prior constraints. However, if we relax the constraint on prior information, reasonable estimation is still possible. This necessitates another form of estimation where such prior information is directly usuable. In this project, this takes the form of Bayes' theorem.

2.4 Bayesian Estimation

The crux of this form of estimation is to infer the goal, the density function, from given measurement data by using Bayes' theorem as a framework. Bayes' theorem delivers a probabilistic function that describes our degree of belief in our estimation. This belief can be maximised to obtain the best prediction based on the evidence given to the estimator. This introduces the potential for adding other experimental data to allow the estimator to become more accurate.

2.4.1 Bayes' Theorem

Bayes' theorem here takes the form of a function that describes the probability of a density function (or any general belief) such that [21]:

$$p(f|m) = \frac{p(m|f)p(f)}{p(m)} = \frac{p(m|f)p(f)}{\int_{-\infty}^{\infty} p(m|f)p(f)df} \tag{2.34}$$

There are three crucial contributing factors in (2.34) that quantify the belief in a density function given some measured data. These are:

1. the prior, $p(f)$, describes the initial degree of belief regarding the density function f,
2. the posterior, $p(f|m)$, describes the belief of a proposed density function f given certain measurements m, and
3. the quotient, $\frac{p(m|f)}{p(m)}$, which models the belief the measured data m gives for a certain hypothesised density function f.

2.4.2 Multivariate Gaussian Distributions

To establish the basis of our Bayesian analysis for the design in Chapter 3, the multivariate Gaussian distribution will be used to describe the probabilities in (2.34).

If we were to have a random function discretised into a vector f (where $f \in \mathbb{R}^{N_y}$) with Gaussian variation, then its distribution would be described by [21]:

$$f \sim \mathcal{N}(\mu_f, C_f) \Leftrightarrow p(f) = \frac{1}{(2\pi)^{N/2}\det(C_f)^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(f-\mu_f)^{\intercal}C_f^{-1}(f-\mu_f)\right) \tag{2.35}$$

The covariance matrix ($C_f \in \mathbb{R}^{N_y \times N_y}$) in (2.35) describes the variance of each component of the vector and the covariance of each point with each other point. This describes the dependence of a data point to its neighbouring points. This becomes an essential part of the design in Chapter 3.

2.4.3 Bayes' Theorem for Multivariate Gaussian Distributions

Combining the theory of Bayes' theorem and multivariate Gaussian distributions, we can obtain an analytically tractable statement estimating a random vector with respect to a prior vector. If we have two multivariate random variables x and y, then if

$$x \sim \mathcal{N}(\mu_x, C_x), \text{ and} \tag{2.36}$$

$$y|x \sim \mathcal{N}(Mx + \mu_y, C_y) \tag{2.37}$$

then according to Bayes' theorem, the posterior is [22]:

$$x|y \sim \mathcal{N}(R(y - M\mu_x - \mu_y) + \mu_x, C_x - RMC_x{}^T), \text{ where} \quad R = C_xM^T(MC_xM^T + C_y)^{-1} \tag{2.38}$$

This derivation is crucial for the design of the density function Bayesian estimator in Chapter 3. This expression directly combines multivariate Gaussian distributions and Bayes' theorem into a usable format.

2.5 Background Conclusions

Throughout this chapter, we have explored the physical implications of NMR T2 relaxation, especially for bound fluid fraction in a porous media. In addition, the current literature on inversion techniques have been examined, and their implications have been analysed. Finally, we have established the foundational theoretical background of Bayesian estimation. This technical background guides the design of the estimator in Chapter 3.

Chapter 3

Design

The design of the estimator was in direct correspondence with the system it describes and predicts. With the Bayesian framework, we obtain the belief of an estimate with respect to specially prepared prior information. There are several assumptions made to allow the estimator to be simple yet robust for application.

Figure 3.1 describes the overall computation workflow of BFF estimation. The parts of this system that will have specific attention towards it are: preparation of the prior information, the estimator, integral transform to obtain the BFF, and how the error is analysed.

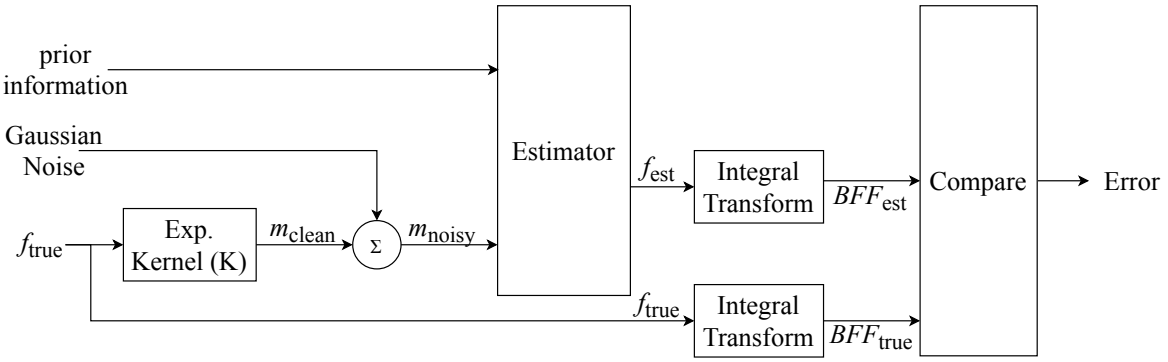


Figure 3.1: The topology of the test architecture of the proposed BFF estimator

3.1 The Bayesian Model

The derivation of the analytic expression of the Bayesian estimate density function requires a series of modelling choices so that we may satisfy the following goals:

1. full integration of a precomputed density function prior into the expression,
2. the estimate density function is constrained to be non-negative,
3. there is design flexibility for assigning uncertainty to different measurements, and
4. the estimation process uses only the measurement data, m, kernel, K, measurement noise variance σ_ϵ, and the prior.

The discussion of the creation of this model expands on work done by Teal [23]. The topology of the framework is shown in Figure 3.2. We seek to utilise Bayes' theorem from (2.34) and (2.38). This yields us an analytic expression describing the T2 density function conditioned on the measurement data, $p(m|f)$,and expected density functions, $p(f)$.

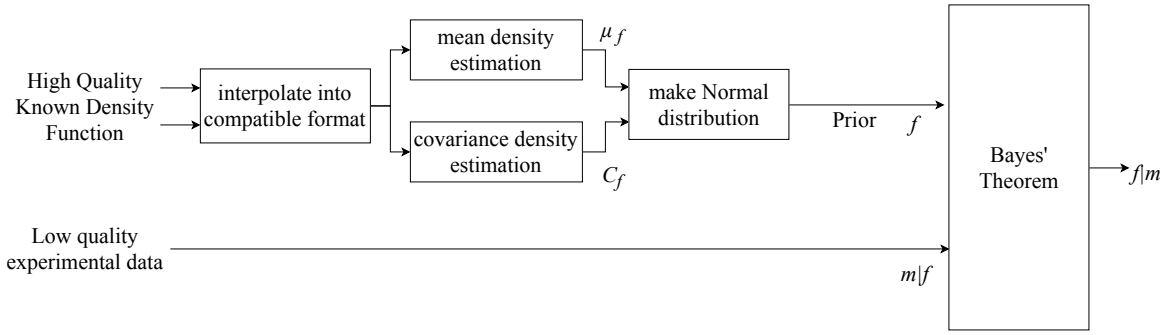


Figure 3.2: Structure of the Bayesian framework of the estimator

3.1.1 Gaussian Assumption for the Prior $p(f)$

To allow for tractable manipulation of the probability density functions, we assume each of the discretised points that make the T2 density function have a Gaussian distribution:

$$f \sim \mathcal{N}(\mu_f | C_f) \tag{3.1}$$

where μ_f is the mean expected prior density function vector that we would expect and C_f is the covariance between the different points of f.

There are three major implications with this assumption:

1. A Gaussian function's domain allows for a non-zero probability for negative values, **domain** $\mathcal{N} \in (-\infty, \infty)$, meaning there is a possibility of a negative density function. This would violate the non-negativity constraint of the model.

2. A Gaussian distribution may not well describe the true distribution of each T2 relaxation bin. In practice this was found not to be an issue

3. The prior may not be accurate if we have a small number of density functions making it up.

The first implication complicates the correctness of the model, as there is a danger of violating the physical non-negativity constraint. In order to provide an analytically tractable result, we make a trade off and relax this constraint. The prior mean is still non-negative as it corresponds directly to physical measurements, but an instantiation drawn from this prior will not have this constraint.

The third implication means that the prior must be constructed from a reasonable sample size of relevant T2 distributions. Therefore, this constrains usage to situations where we already know the type of density function we are trying to find. This is valid for estimating the BFF in a noisy environment where we cannot reasonably expect a controlled laboratory quality conditions.

3.1.2 Modelling the Measurement $p(m|f)$

We model the measured data conditioned on the density function with:

$$m|f \sim \mathcal{N}(Kf, C_m) \tag{3.2}$$

The expression in (3.2) assumes that there is no offset on the measurement data ($\mu_m = 0$). The C_m matrix describes the covariance of each of the measurement points. This allows

for flexibility in adapting for non-uniform SNR measurement data and known dependence between measurement points.

The assumption of a Gaussian distribution of the measurement noise is the same as that made by previous publications as detailed in Section 2.2.1.

3.1.3 Expression for the posterior $p(f|m)$

Now that we have expressions for the measurement and the prior density function, we can now form the expression for the estimate density function conditioned on the measurement data. Adapting (2.38) to (3.2) and (3.1), we get our posterior expression of:

$$f|m \sim \mathcal{N}(R(m - K\mu_f) + \mu_f, C_f - RKC_f{}^T) \text{ , where } \quad R = C_f K^T (KC_f K^T + C_m)^{-1} \tag{3.3}$$

With this analytic expression we may determine our mean posterior density function estimate. In this case we have the same constraints on prior information as the other published techniques *except* for the prior density function. Therefore, design of this additional prior is necessary for the estimator framework.

3.2 Estimator Assessment With Cross Validation

The introduction of highly descriptive priors necessitates a test evaluation method that penalises less generalised estimators. This is where leave one-out cross validation is utilised [24]. The process is as follows:

1. Take one experimental sample and simulate noisy measurement data from it, giving the measurement, m, the input for the estimator.
2. Take the other experimental samples and form the covariance, C_f, and mean prior μ_f from them.
3. Use the posterior expression in (3.3) to form the Bayesian estimate of the density function from the noisy measurement data.
4. Evaluate the bound fluid fraction or other integral transform from this estimate density function.
5. Evaluate the error between the true value and the estimate.
6. Repeat steps 1 to 5 for all of the experimental samples. Average them out to quantify the general performance of the estimator.

This exhaustive technique allows for evaluating the usability of the estimator for general use. The simulated measurement's true density function is unavailable to the estimator so that we may consider general performance. The choice of this evaluation method of the prior also forms the crux of the evaluation in Chapter 5.

3.3 Construction of the Prior

The most essential component of the estimator is the combination of prior high quality experimental data into (3.3). The development of this aspect is based on the Gaussian prior distribution of $p(f)$ is described in Section 3.1.1.

The multivariate Gaussian of the density function prior requires three aspects to be designed such that it can work as intended. These are:

1. interpolation of high quality experimental data into a format compatible with the framework,
2. creation of the mean (μ_f) of all of the high quality representative density functions, and
3. creation of the covariance for all of the T2 relaxation bins (C_f).

High quality experimental data forms the basis for the prior in the Bayesian framework. Thirty NMR T2 relaxation experimental density functions obtained from Schlumberger Doll Research form the prior [25]. These high quality measurements reflect true rock data that make the technique's performance representative of typical application and use.

3.3.1 One Dimensional Interpolation

The prior estimate must be compatible with the rest of the Bayesian framework to be usable. For example, 100 T2 relaxation bins may describe a high quality experimental T2 density function. However, it may requires conversion into 30 relaxation bins to be compatible with the estimator framework. Interpolation of the prior to the actual framework's dimensionality is used to bridge between these two domains.

Any extrapolation is set to zero as we assume that the measurement tool is only sensitive for the range of T2 relaxation values of the T2 axis we provide.

Figure 3.3 demonstrates different interpolation schemes for a conversion from 10 T2 relaxation bins to 100 T2 relaxation bins. Examining all of the candidate interpolation techniques, we can see the following:

1. Nearest point and next point interpolation maintains the discretised coarseness of the original input function, making them poor at providing a representative T2 density function.
2. The linear interpolation is less coarse than the nearest and next point interpolation techniques but it does not provide a smooth function. This makes it insufficient for providing a function with the expected smoothness of a density function.
3. The spline interpolation [26] and Makima (Modified Akima cubic Hermite) interpolation techniques [27] directly violate the non-negativity constraint on the density function.
4. This leaves the shape-preserving piece-wise cubic interpolation (PCHIP) technique [28] as the best candidate. It returns a valid non-negative function that is smooth. These two crucial aspects make it representative of a T2 density function attained through NMR relaxation.

The resulting design decision made was to use PCHIP interpolation. It is fit-for-purpose in adapting experimental data to the Bayesian framework while preserving properties that accurately describe T2 relaxation.

3.3.2 Estimation of the Prior Mean μ_f

Computation of the mean of the prior involves taking the arithmetic mean of all of the experimental density functions over each of the T2 relaxation bins [29]:

$$\mu_f = \frac{1}{N_{\text{rocks}}} \sum_{i=1}^{N_{\text{rocks}}} f_i \quad \text{, where} \quad \mu_f \in \mathbb{R}^{N_y} \tag{3.4}$$

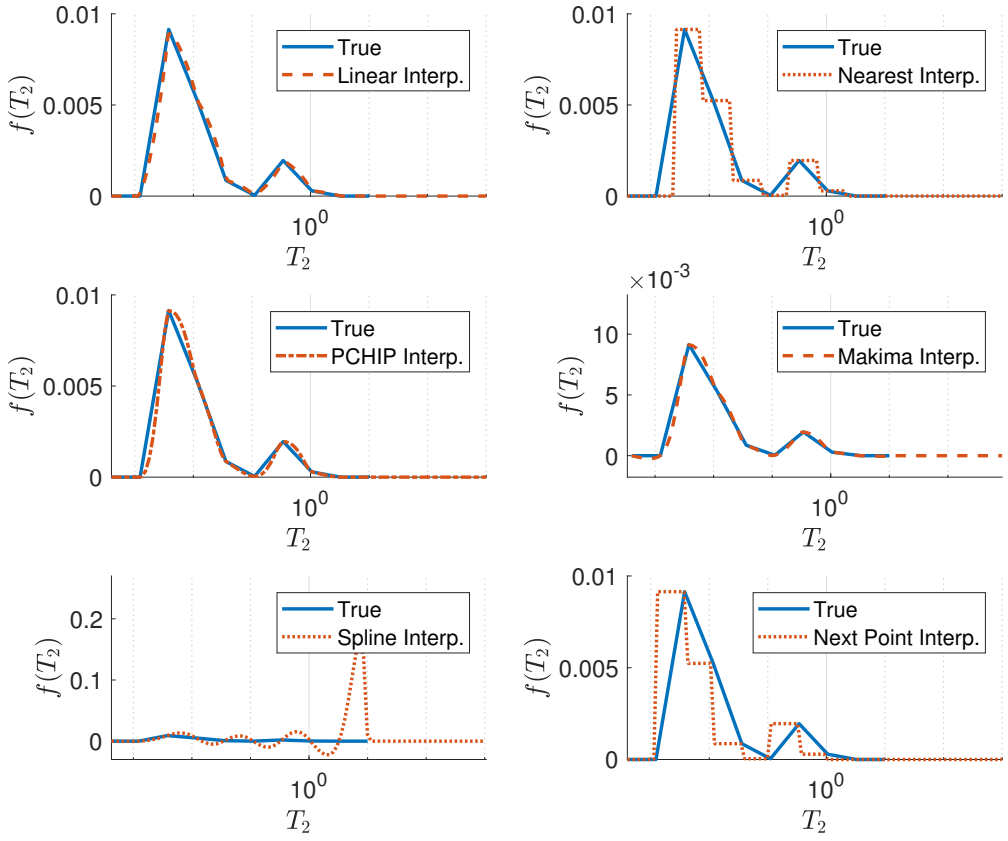


Figure 3.3: Test of different interpolation techniques for interpolation from a lower dimensionality ($N_y = 10$) to a higher dimensionality ($N_y = 100$)

This is the estimation of the expectation of all of the T2 density functions of porous media that make up the experimental data.

3.3.3 Estimation of the Covariance C_f

The design choice of choosing the covariance estimator of the prior density function is concerned with balancing between measurement data and the prior information. Furthermore, there may be a dependence between each of the T2 relaxation bins that may be exploited to allow for a more robust estimator. Here we will analyse three candidate estimates for the covariance. These are that the prior covariance, C_f, is either:

1. uniform and independent,
2. non-uniform and independent, or
3. non-uniform and dependent.

Method A – Uniform Independent Covariance Estimation

The first candidate of the estimation of the covariance assumes that the uncertainty for each T2 relaxation bin is uniform and independent. This makes the covariance equivalent to a scalar multiplied by an identity matrix:

$$C_f = \sigma_\epsilon \cdot I \tag{3.5}$$

This has the benefit of high generality as we can set a blanket uncertainty for our prior measurements. However, this also reduces the descriptiveness of the prior.

Method B – Non-Uniform Independent Covariance Estimation

This estimation of the covariance of the prior density function takes into account the varying uncertainty for different T2 relaxation bins.

$$C_f = \begin{bmatrix} \sigma_{f_1} \\ \sigma_{f_2} \\ \vdots \\ \sigma_{f_{N_2}} \end{bmatrix}^T \cdot I, \text{ where } \quad \sigma_{f_i} = \sqrt{\frac{1}{N_{\text{rocks}} - 1} \sum_{j=1}^{N_{\text{rocks}}} (f_{i,j} - \mu_i)^2} \tag{3.6}$$

The assumption of independence leads to a covariance that is a positive definite diagonal matrix (all of the non-zero values are positive and on the diagonal). Though more flexible than method A, this assumes that the high quality experimental T2 relaxation bins do not have dependence on one another.

Method C – Non-Uniform Dependent Covariance Estimation

This estimates the covariance directly from the high quality density functions [29].

$$C_f = \frac{1}{N_{\text{rocks}}} \sum_{i=1}^{N_{\text{rocks}}} (f_i - \mu_f)(f_i - \mu_f)^T \tag{3.7}$$

This models the dependence between different T2 relaxation bins in the density function. Hence, this covariance estimate fully describes the uncertainty and dependence of different points of the prior. However, it is more vulnerable to overfitting directly to the prior data we are using.

Choosing the Covariance Estimate

To make our choice of the covariance of the prior, we must determine which one tends to perform the best. The methodology of the experiment is as follows:

1. Scale the covariance with a scalar, α, such that we may see how the candidate over estimates and underestimates the true uncertainty. This is analogous to the regularisation in the approximate ILT in Section 2.3.2.

2. Utilise leave-one-out cross validation (detailed in Section 3.2) so that the true density function we are estimating is unavailable to the estimator itself. This is so we may penalise overly specialised prior distributions. We average out all of the results with root mean square error so we may penalise estimation bias and variance equally [30].

3. Create an estimate of the BFV from the density function estimated by the posterior expression in (3.3) and compare with the true BFV. We operate at $T_c = 33$ ms as this will be the typical operating point for the estimator [31]. The BFV prediction uses the sharp integral transform (3.8) given in Section 3.4.

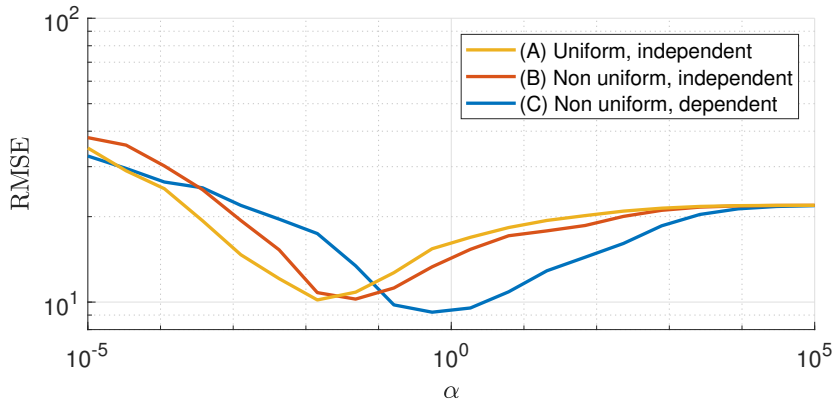


Figure 3.4: Comparison of the root mean square error of the estimator for the uniformly scaled candidate covariance estimates, αC_f

Figure 3.4 illustrates the performance of the candidate estimates of the covariance at an SNR of 10. We see that the non-uniform and dependent covariance estimate (Method C) outperforms the other estimates for where $\alpha > 0.1$. Methods A and B at their best respective α are outperformed by method C. As we are looking for the best possible covariance prior, the non-uniform dependent covariance (Method C) is the most ideal. In the final design we utilise $\alpha = 1$ as it still outperforms the other estimates while simplifying the model with little cost.

3.4 Integral Transform

The goal of the estimator is to estimate the bound fluid fraction of a porous media sample, not the density function. Obtaining this value requires the computation of integral transforms of the estimated density function for the bound fluid volume and the porosity. The two types of BFV transforms are shown in Figure 3.5 for where $T_c = 33$ ms. Both candidate transforms are analysed in the comparative evaluation in Chapter 5.

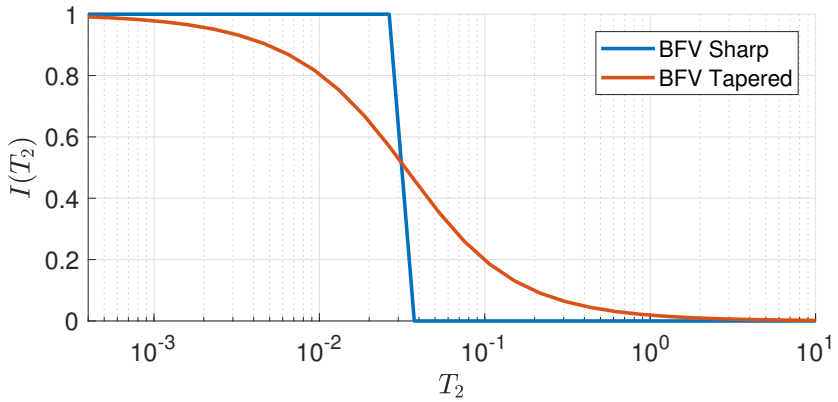


Figure 3.5: Comparison of the BFV integral transforms for a bound fluid cut-off of $T_c = 33$ms

3.4.1 Sharp Bound Fluid Volume

The bound fluid volume is defined as the integration of the density function from $T_2 = 0$ to $T_2 = T_c$ (2.11). In the discretised version, this is equivalent to the inner product of the density function with a vector composed of ones for values below Tc and zeros for values above. There is no smooth transition from the bound fluid to the free fluid; hence, it is referred to as the sharp Bayes method. This is expressed as:

$$BFV_{\text{sharp}} = \int_0^{T_c} f(T_2)dT_2 \approx \text{step}(T_2 - T_c)^T f \tag{3.8}$$

The strength of using an integral transform like this is that it is explicit with what it considers bound fluid volume. The weakness of this however is that it does not have any tolerance for 'leaking' from one T2 relaxation bin to another. This is something that a noisy measurement can inflict onto an estimation of the density function. This brings into the picture another candidate; the tapered integral transform.

3.4.2 Tapered Bound Fluid Volume

Tapered integral transforms have a looser tolerance between bound and free fluid volume. A looser tolerance allows for a more robust estimator for noisy environments that may cause 'leakage' between T2 relaxation bins. The candidate integral transform for comparison with its sharp counterpart is the exponential Haar transform (EHT) proposed by Gruber et al. [1]. The discretised version or this is the inner product between a discretised EHT and the density function. This is depicted analytically as:

$$BFV_{\text{tapered}} = \int_0^{\infty} K_{\text{EHT}}(T_2, T_c)f(T_2)dT_2 \approx K_{\text{EHT}}^T f \tag{3.9}$$

3.4.3 Porosity

Calculating the porosity of the discretised density function is more straightforward. It is by definition the sum of all the T2 relaxation bins (2.9). This is an approximation of the integration of the density function over the entire T2 domain.

3.5 Metrics

A vital aspect of estimation is to have defined metrics to evaluate and compare different techniques. The goal in this part of the design is to fairly evaluate the viability and effectiveness of the different techniques in a balanced manner.

3.5.1 Error

The error of the estimated BFF will take the form of an absolute error. A comparative error is used instead of comparing the actual BFF estimate as we are utilising cross validation over thirty different samples of high quality experimental data. We want to establish a generalised measure of performance.

To visualise the spread and bias, we will use the empirical cumulative distribution functions (CDFs). These are especially advantageous as we may directly compare the cumulative probability of the error for different techniques. We may explicitly state the likelihood of a degree of error so that we may see how performance varies for different magnitudes of acceptable error.

3.5.2 Computational Effort

Though not exhaustive, the computation time of the algorithm reveals a useful comparative perspective on feasibility. For an algorithm to be feasible in the field, it should have a similar or lower computation time to other techniques. The computer and its computational load are held constant. Several repeated results are taken to obtain a general distribution of the computation time.

3.6 Design Conclusions

The consideration of the several intricacies of the Bayesian framework has made its development possible. The flexibility of a Gaussian density prior combined with the analytic expression in (2.38) provides a strong case for its implementation. The following chapter explores the implementation of this design so that it may be compared with other techniques in the literature.

Chapter 4

Implementation

The Bayesian estimator proposed in this report has required comprehensive technical implementation to provide a strong comparative analysis of its effectiveness. These components include:

- implementation of techniques proposed by previous studies for this problem domain,
- the validation of the implementation of those competing techniques such that each of their distinct features can be accurately compared and analysed,
- creation of the proposed estimator with consideration for its special input requirements compared to the previous techniques, and
- a fair test procedure and algorithm to compare and contrast the different estimators.

These several components were developed, prototyped, and analysed in *MATLAB*. This is such that the analysis may be general. Therefore, the results from here can be transferred towards other platforms such as an embedded implementation.

4.1 Existing Techniques

The existing techniques were implemented and validated against the published results. This allows for analysis beyond the specific case study density functions discussed within their respective publications. This means that we may obtain a valid comparison of methods via a larger variety of representative data. In addition, there is no known publicly published implementation of the techniques by the authors, hindering implementation.

4.1.1 Validation of Implementation

The validation framework compares an extraction of the original estimate with the reproduced estimate. This allows for an accurate perspective of any analysis made between the benchmark estimators and the proposed estimator. Figure 4.1 demonstrates the validation topology used by this project.

The components of the validation framework process are as follows:

1. The image extractor uses an image of the model in the publication to acquire the published density function and its technique's estimate. An off the shelf image extractor was used [32].

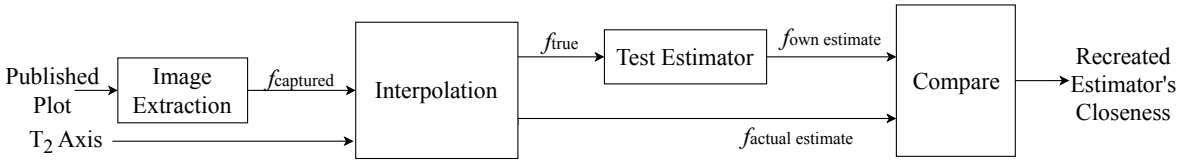
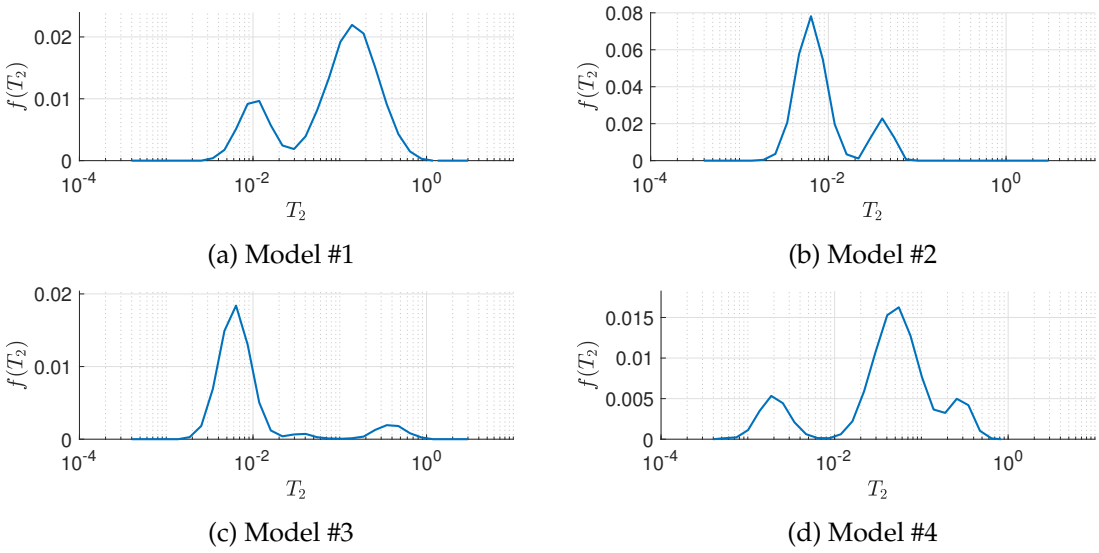


Figure 4.1: Test architecture for recreating the existing estimators in the literature

2. The interpolator discussed in Section 3.3.1 is used so we may have direct compatibility and comparability between the different frameworks
3. The estimator being validated develops its estimate of the density function
4. The comparison component allows for analysis of the deviations between the true estimate and the recreated estimate. This takes the form of visual display and the numerical indicators of error and deviation.

We use the published model density functions analysed by Gruber et al. [19] for the validation of the different recreations of the techniques. This means that any analysis we make on technique accuracy here may be representative of actual data. Figure 4.2 illustrates each of the models used for system validation.



(a) Model #1

(b) Model #2

(c) Model #3

(d) Model #4

Figure 4.2: The T_2 density functions of the models used for system validation

4.1.2 Inverse Laplace Transform Approximation (ILT)

The canonical benchmark for density function estimation is the ILT approximation detailed in [15] and Section 2.3.2. This project explores one-dimensional density functions rather than the two dimensional kind detailed by Venkataramanan et al. [15]. This meant that the re-implementation tended towards the one-dimensional case proposed by the Butler-Reeds-Dawson method instead [13].

Non-uniform SNR

The simulation implementation from Gruber et al. [19] is complicated by the inclusion of additional short pulse sequences that accompany the typical long pulse sequence. There are

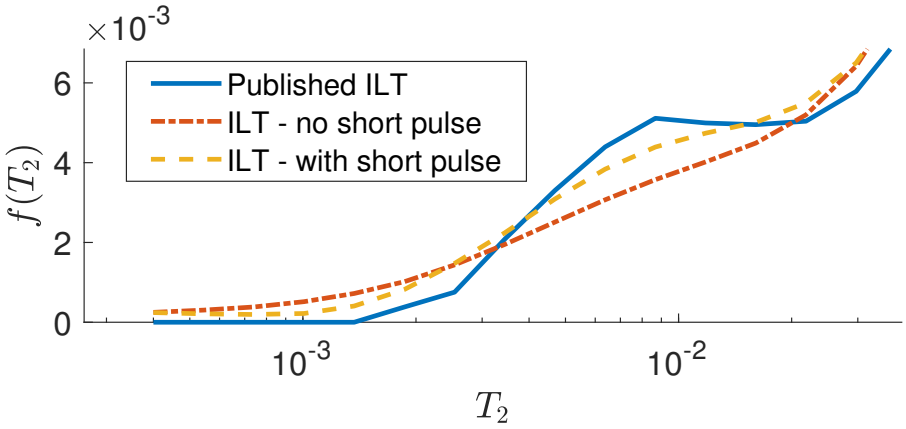


Figure 4.3: Recreation of model 1 (Fig. 4.2a) ILT from Gruber et al. [19] with and without the inclusion of repeated short pulse sequences.

a series of thirty pulses repeated 10 times that increase the SNR for short relaxation times. This means that to reproduce the technique exactly, this component had to be reproduced as well. The method of implementation is as follows:

1. Simulate the data for only 30 samples in the time domain 10 times (short pulses)
2. Average these 10 pulse sequence results together
3. Replace the first 30 values of the long sequence with the new averaged short pulse sequence results
4. Scale the first 30 rows of the kernel matrix, K, by the SNR increase given by the short pulse sequence. We do the same for the measurement vector.

Figure 4.3 demonstrates the effect of accounting for a short pulse sequence in estimator recreation. As we can see, the reproduced technique's estimate is more sensitive for short relaxation. This in turn corresponds to a closer reproduction to the published result. Therefore, this methodology provides a more accurate representation of short relaxation times for our reproduction of the ILT experimental results in [19].

Scaling the Noise

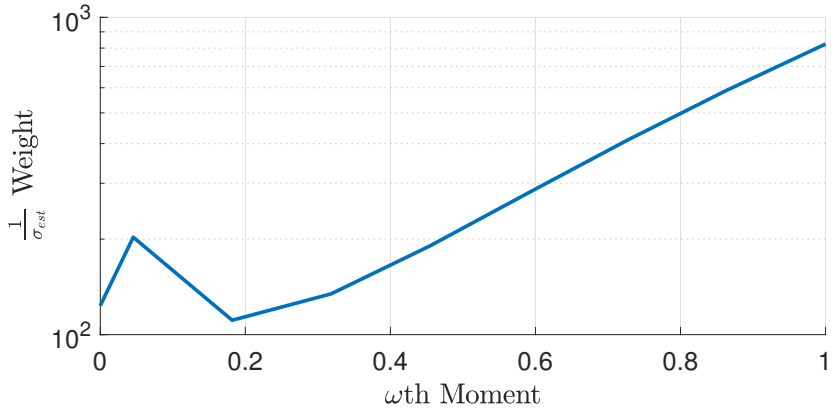
The published noise standard deviation of 0.1 and an SNR of 10 in [19] requires that the signal power be unity. However, the density functions benchmarked in the publication do not have unity power, i.e, the power of the measured signal before noise is added is not unity ($P_{\text{signal}} = \frac{1}{N_2}\sum_1^{N_2}(m_i)^2 \neq 1$). Therefore, the noise standard deviation is scaled to achieve the published SNR such that the results can be recreated.

T_2 Axis

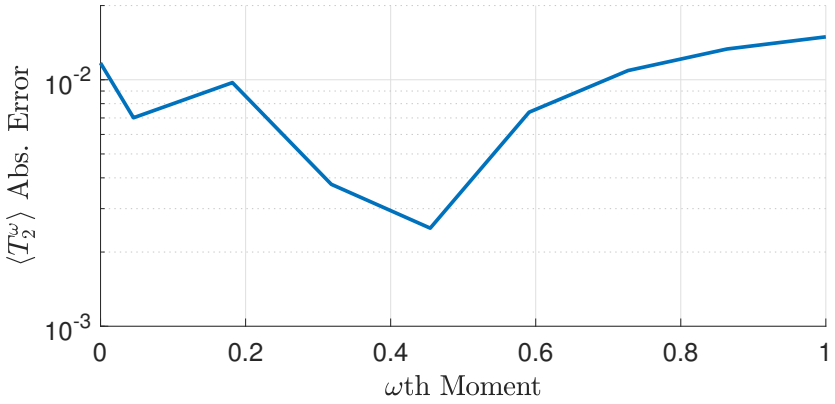
The T_2 axis of the density function estimate is a crucial aspect to reproduce due to the high variability that may come from an ill-conditioned problem. The time axis was assumed to be defined for times from $2 \times t_{\text{sample}}$ to 3 seconds as the simulation results in [19] are bounded within these values. The $2 \times t_{\text{sample}}$ bound is justified by the Shannon Sampling theorem [33]. A sample period of t_{sample} may only fully reconstruct T2 relaxation times of more than $2 \times t_{\text{sample}}$. In addition, the attention to this specific aspect of an inversion is justified by recent research into estimation sensitivity for that region of the T_2 domain [34].

4.1.3 ILT+

This technique requires a significant amount of complexity as it combines moment estimation by Venkataramanan et al. [20], tapered area estimation by Gruber et al. [1], and the initial ILT methods by Venkataramanan et al. [15] and Butler et al [13]. The implementation follows many of the same points as the ILT method discussed in Section 4.1.2.



(a) Weighting of each moment estimation's certainty on a logarithmic scale



(b) Absolute error of the moment estimation on a logarithmic scale

Figure 4.4: A comparison of the certainty alongside the absolute error of the reproduced moment estimate for SNR = 10.

Moment Estimator

The ILT+ implementation was adversely complicated by the inclusion of the moment estimator detailed in Section 2.3.4 and given by Venkataramanan et al [20]. One of the main problems with it was with robustness and error. In order to compensate for this, we modify its weighting so that we may at least acquire indicative results for the evaluation in Chapter 5.

To demonstrate a problem encountered in implementing the moment estimator in the estimator framework, we examine the moment estimation for model 1 from Gruber et al [19] (shown in Figure 4.2a).

Figure 4.4a indicates that as higher moments are estimated, we have a higher weighting for that estimate. However, in Figure 4.4b we can see that the absolute error increases for

where $\omega > 0.5$. This is problematic as the implementation weights these more erroneous points more highly, i.e. the optimisation framework is more certain about worse estimates.

Given this issue with the implemented moment estimator there is a need to diminish its dependence on worse estimates. This diverges from the published results. However, we still preserve indicative trends of performance for the comparative analysis.

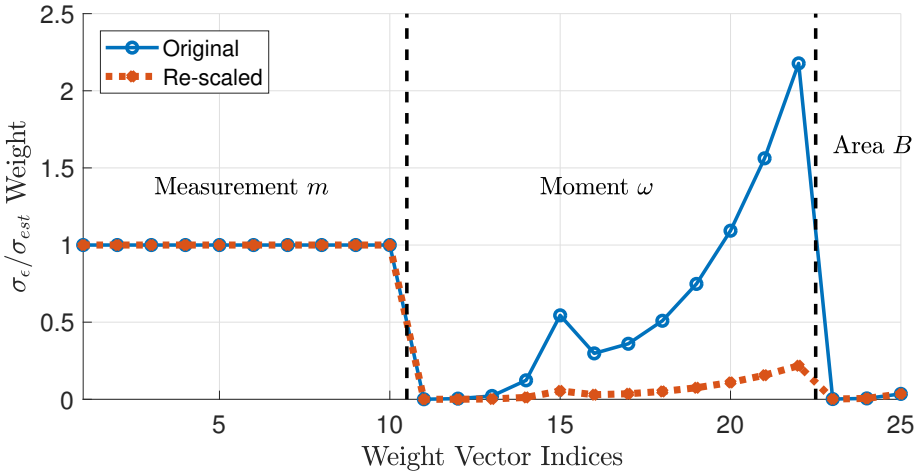


Figure 4.5: Comparison of the original and altered weight vectors for ILT+ to allow for analysis of general indicative performance

Figure 4.5 demonstrates the augmented weighting regime for our implementation of ILT+. The moment weightings are reduced tenfold while all others remain the same so that the moment estimation may not overpower the overall estimation. This amount of reduction was done so that the implementation of the ILT+ would follow to an extent the published ILT+ results.

4.1.4 Tapered Area Estimator

The tapered area estimator has a far simpler set up compared to the previous approximation techniques. This is due to it being fully self-contained and analytically tractable. It contains the exponential Haar transform detailed in [1] and Section 2.3.3. It is a simple dot product between the time domain version of the transform and the measurement data per (2.28)

4.2 Bayesian Technique

The implementation of the Bayesian technique is more comparable to the tapered area estimator rather than the regularised least squares methods of the ILT and ILT+. This is due to the Gaussian assumption of the prior density function and the measurement data permitting an analytic approach to implement the estimator (as discussed in Section 3.1). The implementation is compared with the results obtained by Teal to allow for validation [23].

4.2.1 Estimator Architecture

We adopt a computational sequence for the Bayesian technique to minimise unnecessary computation. The sequence, as shown also in Figure 3.2, is as follows:

1. we take the high quality known density functions and construct our prior mean and covariance before any time sensitive procedures. The prior can be constructed before estimation

2. calculate the estimate density function, the mean of the Bayesian posterior, using (3.3)

The procedure requires no iterative computation like the ILT and ILT+, providing potential for less computation time. This is because the time consuming computation of the prior function does not have to be done within time-constrained use.

4.3 Test Environment

The test environment's implementation required attention so that it provided immediate comparability and fairness. The implementation of the test environment is required to be such that it provides an illuminating and fair comparison of the different techniques. The code for this can be found in [35].

4.3.1 Test Architecture

The test architecture that compares each of estimators aims to demonstrate the uncertainty and typical performance of all of the techniques. In order to keep the Bayesian estimator from directly accessing the exact sample it is testing, leave one-out cross validation (as discussed in Section 3.2) is utilised for the evaluation process. Figure 4.6 details the architecture of the test setup. Table 4.1 outlines the variables set for this test architecture. Any variation from these variables in the evaluation will be explicitly stated.

	Variable	Experimental Value
N_2	Number of Time Samples	2500
N_y	Number of Bins in T2 domain	30
t_E	Sample Period	200 µs
T_2	Axis Bounds	400 µs to 10 s
SNR_{linear}	Signal to Noise Ratio (Uniform)	10
T_c	Bound Fluid Cut-off	33 ms

Table 4.1: The variables used for the test architecture.

An extension of the test architecture can allow for an analysis of how different cut-off times lead to different performance. This provides a perspective of when one technique may be more suitable than another technique. This comes at the cost of computational time, as estimating many samples over N_y different cut off times for several techniques requires a non-trivial amount of computation.

4.3.2 Implementation Conclusions

Throughout Chapter 4, there has been detailed analysis of the implementation of each of the estimators. Most of the complications came from quality issues in reproducing the existing techniques. The absence of published code by the authors made high quality reproduction non-trivial. However, the compromises made still allow for an indication of typical technique performance and trends. These aspects allow for the detailed evaluation in Chapter 5.

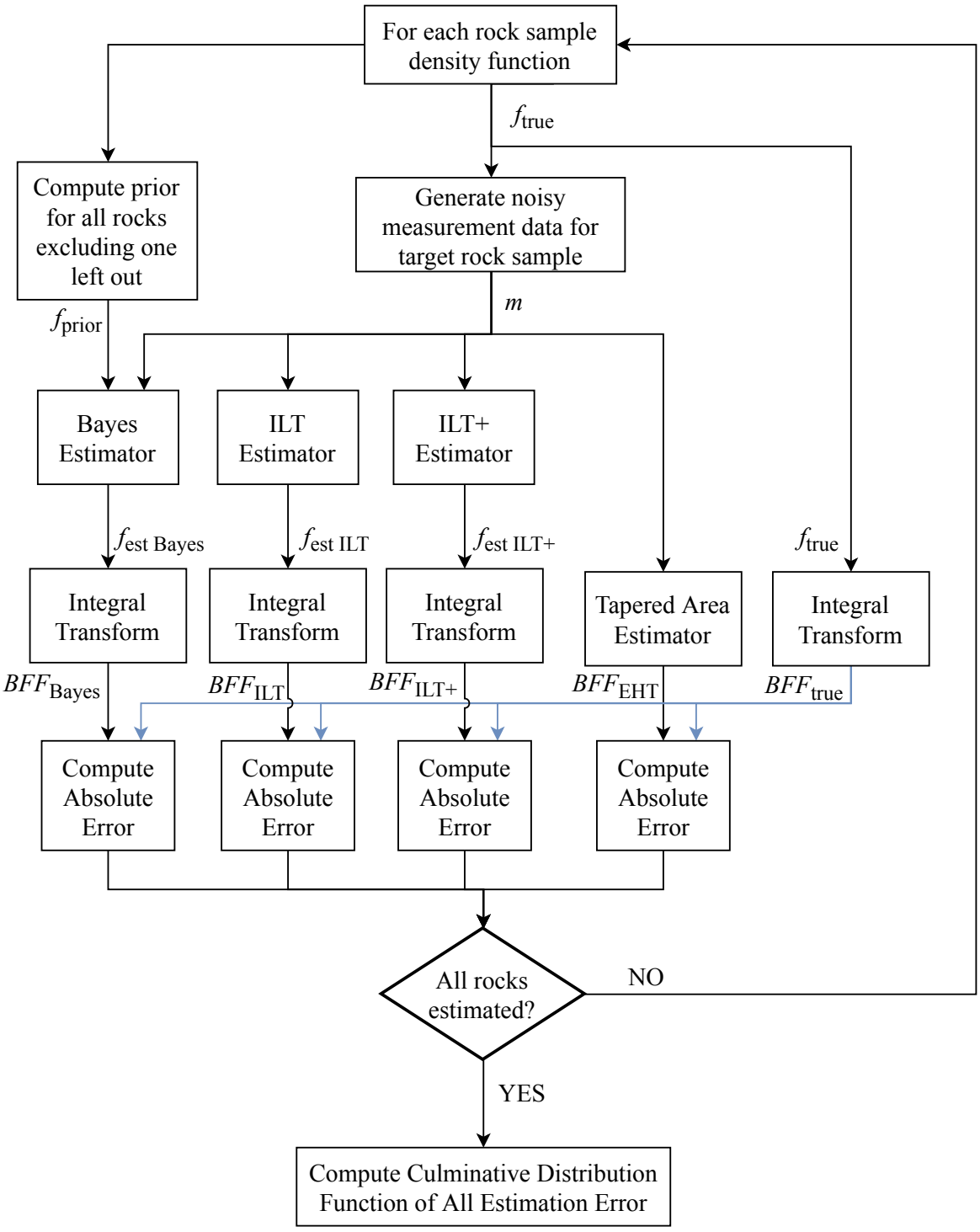


Figure 4.6: The overall test setup to compare each of the estimators fairly.

Chapter 5

Evaluation

Due to the variable nature of statistical estimation, there are several different perspectives that will be explored. Primarily, the Bayesian technique will be compared and benchmarked with the literature's competing methods. This will be analysed in terms of error between the technique's estimate BFF and the true BFF of the sample.

We evaluate the estimators via different perspectives of usability, such as :

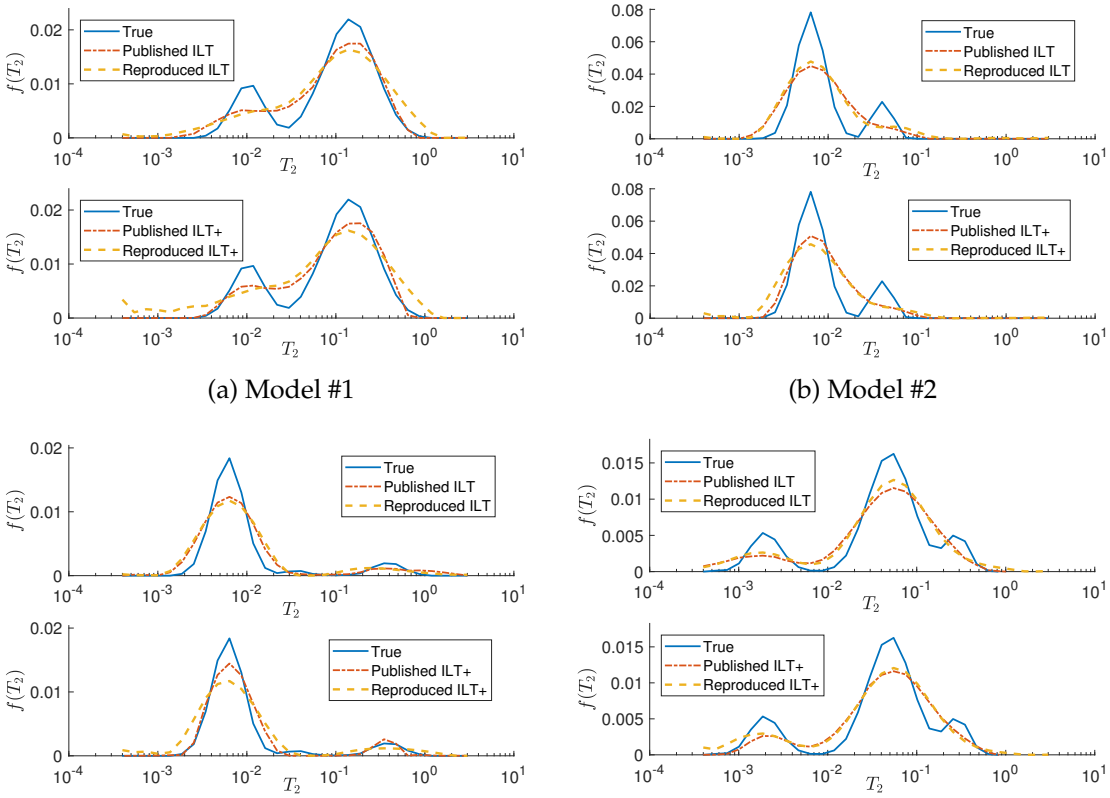
- accuracy at the typical T_c for sandstone rock and at a typical expected SNR environment given by [19], the target operating point of the estimator,
- robustness of the estimation with respect to predicting different bound fluid cut off times,
- sensitivity towards different SNR environments such that we may consider application for different noise powers,
- robustness of the Bayesian estimator to different sample sizes making up the prior, and
- computation speed to indicate potential compatibility with different platforms.

5.1 Recreated Algorithm's Closeness

The analysis in this section utilises the same experimental conditions as detailed in [19]. These results specifically compare algorithm recreation for an SNR of 10. The truncated singular value decomposition compression of the measurement data preserves the 10 largest singular values per [19].

The reproductions of the ILT and ILT+ techniques are compared with published results by Gruber et al [19] in Figure 5.1. The ILT+ technique is particularly erroneous for relaxation times below 3 ms. This is likely to be an error in the implementation of the moment estimator as discussed in Section 4.1.3. Nevertheless, the reproduced techniques, even with error in the replication of the ILT+, still provide some indication of the typical performance of these techniques.

Table 5.1 demonstrates the difference between the overall signal error in implementation and the signal power (porosity of the sample, ϕ) of the overall estimate. We see that the error from estimation tends to be below 10% for the typical density functions. This error does not overwhelm the estimated signal. This means that reproduced performance is indicative of the true application of the algorithm. In addition, if there is a performance gap beyond 10% in the evaluation, it means the implementation error is unlikely to invalidate the performance comparisons made.



(a) Model #1

(b) Model #2

(c) Model #3

(d) Model #4

Figure 5.1: The mean estimated T_2 distribution of the recreated of ILT and ILT+ methods

	Reprod. ILT Error ϕ_E	Published ILT ϕ	ILT Error $\frac{\phi_E}{\phi}$	Reprod. ILT+ Error ϕ_E	Published ILT+ ϕ	ILT+ Error $\frac{\phi_E}{\phi}$
Model 1	1.13E-2	0.1502	7.52%	1.48E-2	0.1507	9.82%
Model 2	1.31E-2	0.2911	4.50%	2.3E-2	0.2854	8.06%
Model 3	9.4752E-4	7.3800E-2	1.28%	9.100E-3	6.94E-2	13.11%
Model 4	5.3E-3	0.1024	5.18%	7.00E-3	0.1013	6.91%

Table 5.1: Error for the mean of the reproduced ILT and ILT+ density function estimates compared to the published results' mean estimate in [19]. The values published here are unit less in correspondence with published results.

5.2 Technique Comparison for a Chosen T_c

The analysis made here is for the typical bound fluid cutoff for sandstone: $T_c = 33$ ms [7]. Figure 5.2 illustrates an empirical cumulative distribution function (CDF) of the absolute error of the estimate bound fluid fraction. The closer the line is to the y axis, the more better it is for a given cumulative probability.

As we can see from Figure 5.2, the Bayesian estimator with the sharp BFV integral transform outperforms all of the other techniques except for at the 99% probability point of the CDF. It has 80% of all estimated BFF below an absolute error of 0.062. Comparatively, the performance of the Bayesian estimator with the tapered BFV integral transform is close to

that of the ILT method. 80% of both the Bayesian tapered and ILT techniques have an absolute estimation error of up to 0.1200. This 50% reduction of error at this cumulative probability indicates a strong case for the viability of the Bayesian framework.

As we move our analysis towards the 95% probability threshold of error, the sharp BFV Bayes and ILT+ methods' performance gap becomes smaller. Specifically, the Bayes Sharp BFV has 95% of all BFF estimations with an absolute error below 0.1295. The reproduced ILT+ in comparison has 95% of its estimations below an absolute error of 0.1479.

The EHT method appear to have undoubtedly the worst performance out of all of the compared techniques. This is likely due to it being the least comprehensive form of estimation. It involves only one integral transform while the other methods include optimisation or a comprehensive prior.

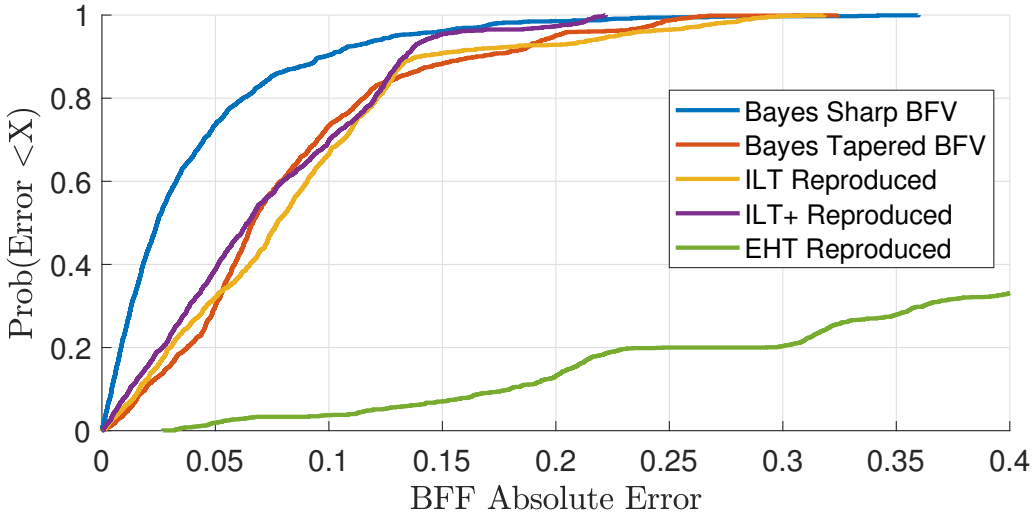


Figure 5.2: Cumulative distribution function of the *BFF* absolute error from the test architecture with respect to the experimental variables given in Table 4.1

5.3 Performance for Different T_c

When we expand the analysis towards a range of different cut-off times for bound fluid fraction, we notice the superior performance of the sharp BFV Bayesian estimator. Figure 5.3 illustrates that the sharp BFV transform performs considerably better than the other techniques except for where T_c lies between 1 ms and 6 ms. At the typical operating points of a bound fluid cut-off of 33 ms (for sandstone) and 90 ms (for carbonates) [3], the sharp Bayes technique far outperforms the other techniques, by up to 0.025 in the estimation error.

The tapered Bayes technique in comparison performs poorly compared to the sharp Bayes technique. This implies that compensating for leaks between the T2 relaxation bins impairs performance. At no point does the tapered method outperform the sharp method. It performs closest to the sharp BFV technique in the short relaxation region of T_c = 2 ms to 4 ms.

The sharp BFV Bayes estimator has an RMSE of less than 0.01 at the bounds of the T2 axis. This indicates that there is little leakage between the density function bins and hence the estimate BFF. This would be due to a near zero prior density function at these T2 relaxation times for all of the experimental density results. The other techniques have a degree of leakage between the relaxation bins that cause non-zero error in their cases. This is an

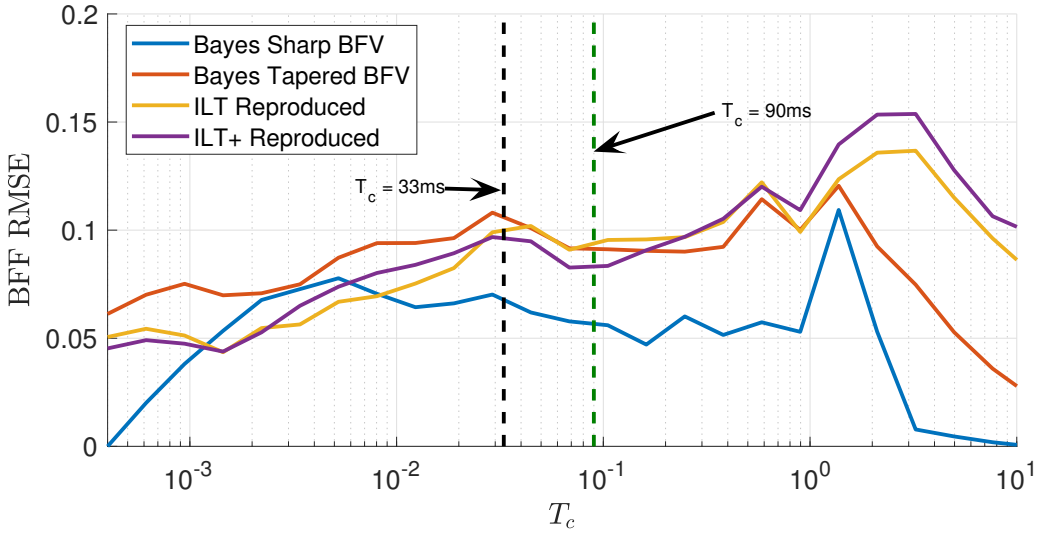


Figure 5.3: Root mean square error for different T_c, SNR = 10. The EHT has been omitted from this analysis to allow closer analysis of the more viable techniques

operating region where the descriptive prior of the Bayesian estimator is at its most advantageous.

5.4 Performance for Different SNR

Estimation of the BFF will be typically done in environments with a potentially variable amount of noise power. This section of the evaluation explores how SNR affects estimation error. Figure 5.4 illustrates an empirical CDF of absolute error for the techniques at different SNR values.

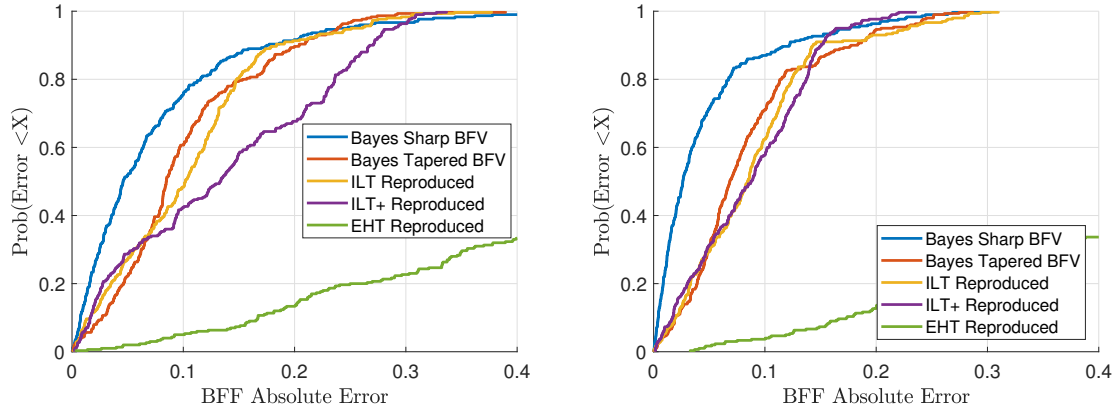
We see that the sharp Bayes method typically performs better than all of the other techniques for each of the different SNRs. For a lower SNR we observe that the sharp Bayes technique tends to perform more similarly to the ILT benchmark. The performance difference becomes more apparent at higher SNR.

As we examine the 80% cumulative probability bound of the error, we see that that as the SNR increases, the sharp BFV Bayes estimator widens the performance gap between the competing techniques. For example, at a poor measurement SNR of 1 the error gap between the ILT and Bayes Sharp BFV is 0.0332. As the SNR increases to 100, this gap more than doubles to an error difference of 0.0766.

We must note that as we examine performance at an increasingly high SNR, we begin to leave the scope the estimator aims to satisfy. This is because high quality experimental results are supposed to be tightly controlled and not influenced by prior information. Detailed prior information may only be suitable in noisier environments where we are unable to control much of the experiment. In those cases, we may accept the use of more detailed prior information to form an estimate.

5.5 Dependence on the Prior

A crucial element for the Bayesian estimator is its robustness to changes in the prior. Figure 5.5 compares the BFF estimate absolute error for different amounts of randomly chosen



(a) SNR = 1

(b) SNR = 5

(c) SNR = 10

(d) SNR = 100

Figure 5.4: Cumulative distribution functions for BFF estimation error for different SNRs

density functions used to construct the prior. We explore the Bayesian estimator with a sharp BFV integral in this analysis. This was repeated via cross-validation for each of the 30 samples 50 times.

This analysis of prior generality illuminates the importance of an effective prior for the Bayesian estimator. As we increase the number of instances that make up the prior, we converge towards a performance bound with diminishing returns. The benefits of an increase to a prior's sample size is pronounced for an increase from 5 to 10 density functions. 50% of the absolute error is below 0.070 for a prior built from 5 instances. The prior constructed with 10 experimental instances has 50% of the absolute error below 0.029. This decrease of error by 58% at this cumulative probability emphasises how a more detailed prior is important for the Bayesian estimator.

The reproduced ILT+, the best of the existing reproduced techniques, tends to generally outperform the Bayesian estimator for where there are 5 or less experimental density functions forming the prior. As the sample size decreases, the Bayesian technique becomes progressively worst as it loses generality. In addition, the Gaussian assumption of the prior density function is less reasonable at these small sample sizes. This exemplifies the Bayesian framework's limitations for generality for where there is little experimental data to form the prior.

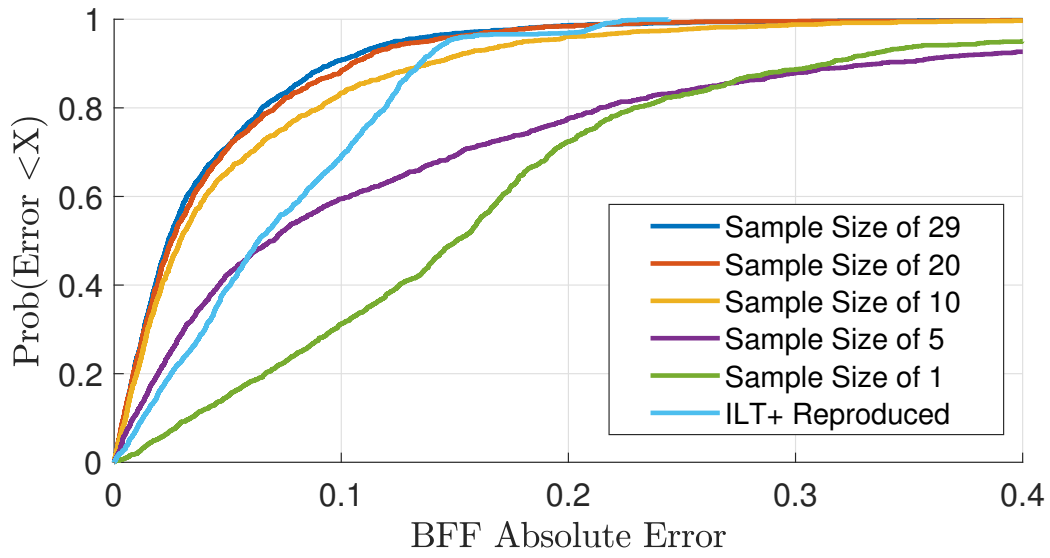


Figure 5.5: CDF comparison of Bayesian estimation absolute error for different numbers of randomly chosen instances forming the prior

5.6 Computation Comparison

Figure 5.6 demonstrates the difference in computation time for each of the techniques. Generally, we see that the proposed Bayesian techniques takes consistently longer than all of the other techniques due to the absence of compressive preprocessing.

The EHT outperforms all of the other techniques, as it requires no matrix inversion or any iterative computation. This comes at the trade off in terms of actual accuracy as demonstrated in Sections 5.2 - 5.4. This means it is served best augmenting another method such as the ILT+ rather than being its own standalone estimator.

The ILT and ILT+ are comparable as they contain the intensive singular value decomposition of (2.25) and multiple matrix inversions of the compressed matrix in (2.23). A matrix inversion has a high computational complexity [36], so if the dimensionality is reduced significantly, performance is decisively increased.

The Bayes method is the slowest as even though there is no iterative computation, it involves the inversion of a term that contains the uncompressed kernel (3.3). This expensive computation implies that compressing the input in some way before introducing it into the framework potentially would yield a better computation speed.

5.7 Evaluation Conclusions

We have determined that given a prior that is representative of what is being estimated, we may estimate the BFF more accurately with the Bayesian estimator. This has been demonstrated for different SNR and different bound volume cut-off times (T_c). However, it is outperformed in terms of computation due to the lack of compression in the framework and is heavily dependent on its prior. Given that experimental data may be available for potential users of this estimator, the latter is less of a concern. The actual performance of the Bayesian technique for these different performance metrics forms a strong case for its viability and use.

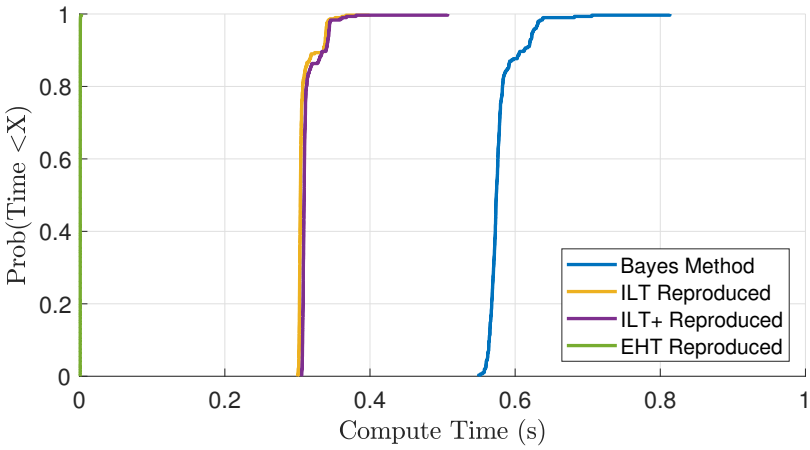


Figure 5.6: Cumulative distribution functions of computation times for each of the techniques

Chapter 6

Conclusions

The project aimed to implement and evaluate the proposed Bayesian estimator alongside the existing published techniques.

In the background, the physical model of NMR relaxation was established. Building off this model, the literature provides a comprehensive benchmark of the current field of estimating T2 relaxation. Satisfying constraints of prior information such as the non-negativity of T2 relaxation times and knowledge of the SNR, they provided robust methods of approximating T2 relaxation density. In nosier environments however, these priori were less useful. An inclusion of a more useful prior grounded in the physical model was required.

The design discusses the how a more useful prior may be implemented with a Bayesian framework. To use a multivariate case of Bayes' theorem such that the it were analytically tractable we assume T2 relaxation timess are to be from a Gaussian distribution. Given this assumption, there is the design question in how we derive the mean and covariance of the prior density function. Allowing for dependence between each of the T2 relaxation bins, at the risk of losing some generality, solves this question. With high quality experimental data embedded into the estimator, we could implement the Bayesian estimator.

Implementation of the existing methods and the proposed method allows for evaluation for different environments such that algorithm robustness and accuracy could be compared. The validation framework built for implementation compares the reproduced and published techniques. The complexity of the ILT+ algorithm complicated direct one-to-one recreation so the reproduction utilised a tuned variation of this. This preserved the indicative performance of the system so that comparison and analysis would be valid. In addition, the proposed Bayesian estimator demonstrated considerable simplicity in implementation in comparison to the majority of the published techniques.

The evaluation of the estimator for different environments in comparison with the reproduced techniques demonstrates the improvements made by using a more comprehensive prior. The Bayesian estimator outperforms the existing reproduced techniques generally for different SNR and bound fluid cut-off. This is done via cross-validation so that we could ensure generality of performance measurement. However, the framework fails if the prior poorly describes what it estimates. With these different perspectives of analysis, we determine that the Bayesian estimator is superior if it has a prior representative of what is measured.

6.1 Future Work

The potential and viability found with the Bayesian framework introduces a potential for its application in other adversely noisy estimation problems. The main parts of this include

adapting the Bayesian framework towards different models.

Specifically, additional potential avenues of future work for the Bayes estimator include:

- Increasing the computational speed of the estimator. There may be potential in adding compressive technologies to get the full benefit of the non-iterative analytic Bayesian expression.
- Examining sensitivity for what forms the prior information, i.e. how well a prior made from only carbonate rock samples may be applicable to sandstone rock samples.
- Can a log-normal noise model be applicable to a tractable Bayesian framework to prevent non-negative density functions.
- Could the Bayes framework augment the ILT+ technique to acquire a general and robust estimator with built in flexibility in prior information.

Bibliography

[1] F. K. Gruber, L. Venkataramanan, T. M. Habashy, and D. E. Freed, "Estimation of petrophysical and fluid properties using integral transforms in nuclear magnetic resonance," *Journal of Magnetic Resonance*, vol. 228, pp. 104 – 115, 2013.

[2] W. Edelstein, G. Glover, C. Hardy, and R. Redington, "The intrinsic signal-to-noise ratio in NMR imaging," *Magnetic resonance in medicine*, vol. 3, no. 4, pp. 604–618, 1986.

[3] M. D. Hürlimann and N. J. Heaton, "NMR Well Logging," in *Mobile NMR and MRI: Developments and Applications* (S. J. V. Edited by Michael L. Johns, Einar O. Fridjonsson and A. Haber, eds.), ch. 2, pp. 11–79, The Royal Society of Chemistry, 2016.

[4] C. L. Epstein and J. Schotland, "The bad truth about Laplace's transform," *SIAM review*, vol. 50, no. 3, pp. 504–520, 2008.

[5] J. G. McWhirter and E. R. Pike, "On the numerical inversion of the Laplace transform and similar Fredholm integral equations of the first kind," *Journal of Physics A: Mathematical and General*, vol. 11, no. 9, p. 1729, 1978.

[6] D. Capitani, A. P. Sobolev, V. Di Tullio, L. Mannina, and N. Proietti, "Portable NMR in food analysis," *Chemical and Biological Technologies in Agriculture*, vol. 4, p. 17, Jul 2017.

[7] W. Kenyon, "Nuclear magnetic resonance as a petrophysical measurement," *Nuclear Geophysics*, vol. 6, no. 2, pp. 153–171, 1992.

[8] Z.-P. Liang and P. c. Lauterbur, "Signal Generation and Detection," in *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective* (M. Akay, ed.), ch. 3, pp. 57–105, Institute of Electrical and Electronics Engineers, Inc, 2000.

[9] R. Akkurt, H. N. Bachman, C. C. Minh, C. Flaum, J. LaVigne, R. Leveridge, R. Carmona, S. Crary, E. Decoster, N. Heaton, *et al.*, "Nuclear magnetic resonance comes out of its shell," *Oilfield Rev*, vol. 20, no. 4, pp. 4–23, 2009.

[10] R. Kleinberg, C. Straley, W. Kenyon, R. Akkurt, and S. Farooqui, "Nuclear magnetic resonance of rocks: T1 vs. T2," in *SPE Annual Technical Conference and Exhibition*, 1993.

[11] R. Kleinberg, "NMR measurement of petrophysical properties," vol. 13, pp. 404–406, 01 2001.

[12] S. Chen, G. Ostroff, and D. T. Georgi, "Improving estimation of NMR log T2 cutoff value with core NMR and capillary pressure measurements," in *SCA International Symposium, The Hague*, pp. 14–16, 1998.

[13] J. P. Butler, J. A. Reeds, and S. V. Dawson, "Estimating Solutions of First Kind Integral Equations with Nonnegative Constraints and Optimal Smoothing," *SIAM Journal on Numerical Analysis*, vol. 18, no. 3, pp. 381–397, 1981.

[14] R. J. Hanson, "A numerical method for solving Fredholm integral equations of the first kind using singular values," *SIAM Journal on Numerical Analysis*, vol. 8, no. 3, pp. 616–622, 1971.

[15] L. Venkataramanan, Y.-Q. Song, and M. D. Hurlimann, "Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1017–1026, May 2002.

[16] L. Eldén, "Algorithms for the regularization of ill-conditioned least squares problems," *BIT Numerical Mathematics*, vol. 17, no. 2, pp. 134–145, 1977.

[17] P. C. Hansen, "The truncated SVD as a method for regularization," *BIT Numerical Mathematics*, vol. 27, no. 4, pp. 534–553, 1987.

[18] R. Kleinberg, A. Boyd, *et al.*, "Tapered cutoffs for magnetic resonance bound water volume," in *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers, 1997.

[19] F. K. Gruber, L. Venkataramanan, T. M. Habashy, P. M. Singer, and D. E. Freed, "A more accurate estimate of T2 distribution from direct analysis of NMR measurements," *Journal of Magnetic Resonance*, vol. 228, pp. 95 – 103, 2013.

[20] L. Venkataramanan, F. K. Gruber, T. M. Habashy, and D. E. Freed, "Mellin transform of CPMG data," *Journal of Magnetic Resonance*, vol. 206, no. 1, pp. 20 – 31, 2010.

[21] C. W. Therrien in *Discrete Random Signals and Statistical Signal Processing*, ch. 2, pp. 31,43, Prentice-Hall, Inc., 1992.

[22] C. M. Bishop, "Pattern Recognition and Machine Learning," ch. 2, p. 93, Springer, 2006.

[23] P. D. Teal, "Bayeisan NMR Relaxometry." `https://gitlab.ecs.vuw.ac.nz/project489-2018/dobbiedavi/engr489Project/blob/master/relevant_docs/pauls_work/notes/bayes3.pdf`, Mar 2018.

[24] M. W. Browne, "Cross-Validation Methods," *Journal of Mathematical Psychology*, vol. 44, no. 1, pp. 108 – 132, 2000.

[25] D. W. Dobbie, "T2distributions.mat." `https://gitlab.ecs.vuw.ac.nz/project489-2018/dobbiedavi/engr489Project/blob/master/test_environment/testing_data/T2Distributions.mat`, October 2018.

[26] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE transactions on acoustics, speech, and signal processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[27] H. Akima, "A new method of interpolation and smooth curve fitting based on local procedures," *Journal of the ACM (JACM)*, vol. 17, no. 4, pp. 589–602, 1970.

[28] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM Journal on Numerical Analysis*, vol. 17, no. 2, pp. 238–246, 1980.

[29] C. W. Therrien, "Discrete Random Signals and Statistical Signal Processing," ch. 6, pp. 301–8, Prentice-Hall, Inc., 1992.

[30] W. Navidi, "Statistics for engineers and scientists," ch. 4, pp. 280–281, McGraw-Hill, 2008.

[31] R. Kleinberg, "Utility of NMR T2 distributions, connection with capillary pressure, clay effect, and determination of the surface relaxivity parameter ρ_2," *Magnetic Resonance Imaging*, vol. 14, no. 7, pp. 761 – 767, 1996. Proceedings of the Third International Meeting on Recent Advances in MR Applications to Porous Media.

[32] A. Rohatgi, "WebPlotDigitizer." `https://automeris.io/WebPlotDigitizer/`, Jan 2018.

[33] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[34] L. Venkataramanan, F. K. Gruber, J. LaVigne, T. M. Habashy, J. G. Iglesias, P. Cohorn, V. Anand, M. A. Rampurawala, V. Jain, N. Heaton, *et al.*, "New method to estimate porosity more accurately from NMR data with short relaxation times," *Petrophysics*, vol. 56, no. 02, pp. 147–157, 2015.

[35] D. W. Dobbie, "test_environment." `https://gitlab.ecs.vuw.ac.nz/project489-2018/dobbiedavi/engr489Project/tree/master/test_environment`, October 2018.

[36] R. Raz, "On the complexity of matrix product," in *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing*, STOC '02, (New York, NY, USA), pp. 144–151, ACM, 2002.

Appendix A

Appendix

A.1 Meeting Notes

Meeting Notes:

Meeting 25

Meeting Date: 18/10/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- finishing final report
- will practice presentation next week w/ Paul

Action points for next meeting:

- full report done
- presentation draft done

Action points achieved since last meeting:

- wrote most of report
- implemented a compromise of ILT+ so that it may be indicative

Action points yet to be achieved since last meeting:

- almost finished report

Meeting 24

Meeting Date: 11/10/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- problem with ILT+ comes from the moment estimator

Action points for next meeting:

- finish report by then
- validate ILT+ as much as possible

Action points achieved since last meeting:

- checked ILT+ validation once again

Action points yet to be achieved since last meeting:

- written conclusion of report 1st draft

Meeting 23

Meeting Date: 4/10/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- got comments back for report sans conclusion

Action points for next meeting:

- write conclusion (3 pages)
- check ILT+ validation, at least near more near to a degree

Action points achieved since last meeting:

- written most of report, gone through first draft process

Action points yet to be achieved since last meeting:

- written conclusion of report 1st draft

Meeting 22

Meeting Date: 27/09/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- setting up test environment, polishing test environment

Action points for next meeting:

- write draft 2nd half

Action points achieved since last meeting:

- write draft 1st half
- sorted out comparison set up for validating recreated models

Action points yet to be achieved since last meeting:

- none

Meeting 21

Meeting Date: 20/09/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- validating ILT+ and ILT. Can get to within R2 = 0.9 for ILT+. Can get close but not completely on.
- can add case by case comparisons with known ILT and ILT+ in Gruber 2013. Only require BFF to be accurate

Action points for next meeting:

- start drafting 2nd half of report
- set up comparison cases for models presented in Gruber 2013. We can make a comparison there.

Action points achieved since last meeting:

- written most of first half first draft, aim to hand in on Friday 21/09
- set up validation environment for ILT and ILT+ methodologies

Action points yet to be achieved since last meeting:

- ILT+ validation yet to be fully 100% completed

Meeting 20

Meeting Date: 06/09/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- possible chance of taking practical T2 relaxation data to further improve on 489 report
- started validation of the ILT and ILT+ techniques for the test environment
- started integration with preliminary report, moving onto writing up design chapter

Action points for next meeting:

- double-triple check ILT+ technique, make evaluation as fair as possible
- Start writing report,
    - hand in Design sections draft 1
- send email to investigate possible experimental data from a

Action points achieved since last meeting:

- in progress to complete checking ILT+ and writing design chapter

Action points yet to be achieved since last meeting:

- most of drafting for the first half of the report has been finished

Meeting 19

Meeting Date: 30/08/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- look into creating Parzan window (mixed Gaussian) rather than simple Gaussian for each data point
- start writing the report, combined portions of prelim into full
- keep in mind feedback for the prelim report when writing

Action points for next meeting:

- double-triple check ILT+ technique, make evaluation as fair as possible
- Start writing report,
    - hand in Intro, Background, Design sections draft 1 (mid-tri wk 2)
    - hand Implementation, Evaluation, Conclusion sections draft 1 (wk7)

Action points achieved since last meeting:

- set up the BBF estimator for:
    4. ILT+ 2013 technique
    5. Tapered area estimator 2013

Action points yet to be achieved since last meeting:

- none

Meeting 18

Meeting Date: 23/08/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- gone over testing system for the BFF estimator
- Could create a Parzan window for combining the prior - a Gaussian mixture

Action points for next meeting:

- set up the BBF estimator for:
    4. ILT+ 2013 technique
    5. Tapered area estimator 2013

Action points achieved since last meeting:

- set up the BFF estimator for:
    1. Bayesian with steep cut off for fluid volume
    2. Bayesian with tapered cut off for fluid volume
    3. Classic 2002 Venk technique
- set up testing environment

Action points yet to be achieved since last meeting:

- set up the BBF estimator for:
    4. ILT+ 2013 technique
    5. Tapered area estimator 2013

Meeting 17

Meeting Date: 16/08/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- tested how Gaussian the prior data points are.
- set up a One-sample Kolmogorov-Smirnov test for each porosity bin in f(T2) - finding that a

collection of points have a Gaussian distribution
- looked into testing the log-normal closeness of the density function prior data points. This process is still pending

Action points for next meeting:

- set up the BBF estimator for:
    1. Bayesian with steep cut off for fluid volume
    2. Bayesian with tapered cut off for fluid volume
    3. Classic 2002 Venk technique
    4. ILT+ 2013 technique
    5. Tapered area estimator 2013
- set up testing environment for the system

Action points achieved since last meeting:

- tested how Gaussian each data point of the prior density function is
- finished off testing different estimate of covariance matrices

Action points yet to be achieved since last meeting:

Meeting 16

Meeting Date: 09/08/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- explored the performance of different estimators for the covariance matrix
- we see that there are diminishing returns at certain estimation of alpha (how much we trust the prior)
- the non diagonal covariance estimator indicates significant performance improvements and less sensitivity for the choice of the alpha (it is around 1, which fits with the hypothesis that there should be equal weighting for that)
- for an analysis of the cut-off time for bound fluid: it would be best served to choose a more arbitrary point so that there can be more useful discussion of the model's effectiveness.

Action points for next meeting (or two):

- test how Gaussian each data point of the prior density function is
    - how true is the model's assumption that the density function is is normal for each point
- create an experiment of the zero uniform offset on the estimated covariance matrix
- set up the BBF estimator for:
    1. Bayesian with steep cut off for fluid volume
    2. Bayesian with tapered cut off for fluid volume
    3. Classic 2002 Venk technique

4. ILT+ 2013 technique
5. Tapered area estimator 2013

Action points achieved since last meeting:

- explored the use of different covariance matrices for the Bayesian estimator.
    - recorded results for 1.3, 1.4, 1.5 version of the Bayesian estimator
- developed a simple alpha predicted with the 2 norm of the covariance matrix

Action points yet to be achieved since last meeting:

- have not made a system that trained an alpha for the estimator - far too high complexity for something that can be simply done with a more suitable covariance matrix

Meeting 15

Meeting Date: 2/08/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- have presented the preliminary presentation. Good feedback given for creating an accessible presentation
- we have a framework where experimental data can be used to form the estimator
    - this uses leave one out cross validation to give an empirical best alpha
- the path for the project is to use different priors formed by experimental data such as
    - covariance of each data point
    - assuming smoothness, the cross-correlation between each data point
- also given feedback on the preliminary report
    - give discussion of results and experiments
    - discuss how algorithms had their recreation validated
    - first letter capital for figures and sections cross referencing (Figure 1, Section 2.1)

Action points for next meeting:

- create a method for finding the best alpha value. Try two methods:
    - create a validation on the training process and choose the optimal empirical alpha value from that
    - try to estimate alpha as we already know C_f and C_n - see if this matches

Action points achieved since last meeting:

- presented the preliminary presentation
- implemented a cross validation system using experimental data to evaluate performance for different alpha values - do not need to normalise data before using it
    - this implementation uses a uniform diagonal covariance for the density function

Action points yet to be achieved since last meeting:

- test how normal the density function experimental data is
- create weighted covariance C_f dependent on previous prior data

Meeting 14

Meeting Date: 26/07/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- was sick for a significant part of this week, could not present
- dealing with a normalisation issue of the experimental data. Normalisation directly assumes the porosity itself.
    - will construct prior with normalised data first so bound fluid fraction can be computed
- could perhaps utilise hyper priors for density estimation

Action points for next meeting:

- normalise and construct the prior
- preliminary presentation in wk 3
- test how normal original data is
- estimate covariance, use it to make prior

Action points achieved since last meeting:

- loaded experimental data onto a script to construct the prior

Action points yet to be achieved since last meeting:

- full construction of a Gaussian prior
- preliminary presentation

Meeting 13

Meeting Date: 18/07/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- constructed the Bayesian technique with a zero mean assume prior density function
- could perhaps rework the density function distribution to be log normal to work around the non-negative constraint on that
- the results gathered use a realistic prior with using the true density function as the mean. This is the best possible circumstance
- will present in front of an ENGR 489 lecture slot for the mid way presentation
- have not received preliminary report results yet

Action points for next meeting:

- construct a prior with high quality experimental data, use for Bayesian technique (assume independence of each data point). This will be compared with synthesised measurement data created from it
- test how Gaussian (or not Gaussian) the experimental density functions are (normality test)
- look into improving the prior of experimental data by creating a realistic covariance (for correlation between points) (use autocorrelation method and covariance method)

Action points achieved since last meeting:

- implemented and evaluated Bayesian technique with a zero mean prior density function

Action points yet to be achieved since last meeting:

- create trimester 2 presentation

Meeting 12

Meeting Date: 28/06/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- Paul will not be able to meet for next two Thursdays (5/7/18 and 12/7/18)
- Was unable to meet up due to exams - compensated expectation as a result
- discussion about MSE in evaluating estimators

Action points for next meeting:

- Start Paul’s Bayesian technique - construct and evaluate it
- Sort out T2 week one presentation

Action points achieved since last meeting:

- Mostly finished Venkataramnan et al Porosity Estimation (April 2015)
    - understanding of how the method works as well as its limitations

Action points yet to be achieved since last meeting:

- none

Meeting 11

Meeting Date: 7/06/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- generate the bias error from individual deltas = f(T2) for all of T2

Action points for next meeting:

- Finished Venkataramnan et al Porosity Estimation (April 2015)

Action points achieved since last meeting:

- progress on Venkataramnan et al Porosity Estimation (April 2015)
- final tweaking of preliminary report completed

Action points yet to be achieved since last meeting:

- none

Meeting 10

Meeting Date: 31/05/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

Action points for next meeting:

- Tweak preliminary report, more detail on the future plan and proposed technique
    - rearrange for work done
    - add more questions to the marker
- Start Venkataramnan et al Porosity Estimation (April 2015)

Action points achieved since last meeting:

- completed full draft

Action points yet to be achieved since last meeting:

- none

Meeting 9

Meeting Date: 24/05/18

Who was present:

- Paul Teal
- Robin Dykstra
- David Dobbie

Meeting Comments/Notes:

- first half draft returned, needs restructuring and rewriting
- requires specific referencing

Action points for next meeting:

- finished first full draft on Monday - (fully use feedback given to improve it)

Action points achieved since last meeting:

- Finished T2 ILT+ Gruber 2013
- Fixed figures in report
- finished first draft of other half of report, ready to combine together into effective script

Action points yet to be achieved since last meeting:

- None

Meeting 8

Meeting Date: 17/05/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- acquired Hurlimann chapter on oil well logging
- format report so that:
    - smallest text is the size of a sub caption

    - remove titles on figures
    - refer to subplots by letters, allows for elegant discussion
    - implement hierarachal writing structure - each chapter and section have a description at the start describing what they will entail.
- Robin will come into next meeting to help out with writing the background and motivation on the NMR material.

Action points for next meeting:

- write up introduction, future work and ILT+ method of preliminary report
- finish debugging Gruber 2013, paper 4
- change formatting of report, make it better

Action points achieved since last meeting:

- Wrote 80% of background of the preliminary report
- Created script evaluating tapered area and moment estimations with the true values
- created script framework implementing ILT+ and ILT together

Action points yet to be achieved since last meeting:

- ILT+ not working yet, need to fix the weighting matrix - make it work as intended

Meeting 7

Meeting Date: 10/05/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- emphasis on images and diagrams for communicating concepts
- place evaluation in terms of mean squared error - easy to keep track of
- ask for Haliburtion Oil Well Logging book from Robin for useful images
- bring USB for Martin Hurlimann chapter from Paul with useful images

Action points for next meeting:

- Half of preliminary report complete - the introduction and background survey
- Finish Gruber Paper 4
- Gather accuracy of tapered areas estimations (for report)

Action points achieved since last meeting:

- Finished Gruber Paper 3 (Tapered Areas)

Action points Yet to be achieved since last meeting:

- none

Meeting 6

Meeting Date: 03/05/2018

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- Given CPMG measurement data - not too sure what method was used to derive the T2 distributions from it
- Paper 3 gives tapered areas and moments to establish the constraints of paper 4 - Gruber T2 distribution
- A major part of this project is to establish Tc - the fraction cut off point between the bound and free fluids in the sample. Paper 3 discusses this
    - good to have kernel made specially to differentiate Tc cut off point
- can ignore the incomplete polarisation (G(t)) modelling in S2.2 of paper 3, want to constrain scope to fully polarised measurements (M(t)).
- do not need to have complete readings of background material due to constraints on project execution. Very direct implementation

Action points for next meeting:

- Complete Gruber 2013 paper 3
- Start Gruber 2013 paper 4 (optimistic goal to complete it)

Action points achieved since last meeting:

- written up summaries of venk 2002 and venk 2010
- started gruber 2013 on linear functionals.

Action points Yet to be achieved since last meeting:

- none

Meeting 5

Meeting Date:

26/04/18

Who was present:

- David Dobbie
- Paul Teal

Meeting Comments/Notes:

- Stopping work on Mellin Transform - got satisfactory results for omega in [0,1]
- making sure documentation for past papers are good - able to revise upon on final report
- moving onto twin papers by Gruber 2013 that builds on this

Action points for next meeting:

- Write up summary of venk 2002 and venk 2010
- Start Gruber 2013 - Estimation of Petrophysical and fluid properties using integral transforms in nuclear magnetic resonance

Action points achieved since last meeting:

- Finished implementing mellin transform venk 2010, written documentation on the process.

Action points Yet to be achieved since last meeting:

Meeting 4

Meeting Date:

19/04/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- look into using logarithms to get around computational errors and problems around float computation problems

Action points for next meeting:

- Bug check implementation of Mellin transform Venk.2010, have pictures and results of bug testing (bug plots of G(\omega) and \delta_i) (before meeting)

Action points achieved since last meeting:

- progress on implementation of Venk. 2010

Action points yet to be achieved since last meeting:

- Not yet completed implementation of Venk. 2010 Mellin transform

Meeting 3

Meeting Date:

13/04/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- Would in future look into implementing the flint code, play around with it to make it work appropriately
- important to recreate the paper's properly to make them comparable
- Can utilise "engauge" to convert image data to data points to recreate the paper's made
- Further on with Bayes, about constructing the Priori for the Bayesian analysis

Action points for next meeting:

- Work onto implementing Mellin Transform of CPMG data (Journal of Magnetic Resonance 206 (2010) 20-31) (note that this is more of a stepping stone paper)

Action points achieved since last meeting:

- finished 1D implementation of Venk. 2002
- Progress on 2D implementation of Venk. 2002

Action points yet to be achieved since last meeting:

- none

Meeting 2

Meeting Date:

06/04/18

Who was present:

- Paul Teal
- David Dobbie

Meeting Comments/Notes:

- got code to guide implementation of 1D example - aiming to implement K1 kernel properly, etc.
- Figured out aim (ambitious) to code up competing algorithms and understand by preliminary report hand in.
- given series of papers for next few weeks of work, get a comprehensive picture of the project.
- Discussion on introductory Bayes, will be focus later on in the project.

Action points for next meeting:

- Finish 1D example of extracting exponential time constants in paper 1 venkataramanan-2002-tensor-fredholm
- Get strong understand of 2D example of it, implement it

Action points achieved since last meeting:

- Read paper (Vol. 50, no 5. May 2002 IEEE transactions on Signal Processing (1053-587x(02)03282-8) venkataramanan-2002-tensor-fredholm.pdf
- Worked on implementing the algorithm in it, not complete

Action points Yet to be achieved since last meeting:

- Implement 1D example fully in paper venkataramanan-2002-tensor-fredholm

Meeting 1

Meeting Date:

20/03/18

Who was present:

- Paul Teal
- Robin Dykstra
- David Dobbie

Meeting Comments/Notes:

- Going over project knowledge base
- With the project, techniques required to compensate for the very poor SNR in underground oil detection with how NMR used
- A magnetic field when constant results in a constant spin characteristic of the protons itself (\omega = \beta\gamma)
- When sending a pulse to redirect the spin (90 degrees for example), we have the regrow to normal spin T1 and the much faster decay in T2.
- We have T2* for a non-constant magnetic field, dealing with decay caused by a non-uniform magnetic field since the 2D vectors drift until they oppose.
- To get around this faster decay, we send a 180 degree pulse of RF radiation to reorient it so that it acts in the envelope of the T2 decay curve – we get discretised measurements of T2. (these are called echos)
- When a molecule is closer to the surface, for a fluid (bound) rather than deeper inside (free), it will decay faster
- We can get different t2s for different protons and their different physical positioning
- If we take several measurements of T2 exponential decays we can make a statistical reference of this data and see the prominence of both (we have sum of exponentials with different modifiers on T2)
- Looking at either T2 or T1 combined gives a 2.5D representation of it, this is in the paper Vol. 50, no 5. May 2002 IEEE transactions on Signal Processing (1053-587x(02)03282-8)

- This paper utilised solving for the distribution from discretised data with the Kroncher product
- Note that SNR is defined differently in this paper than its conventional meaning.
- Initially, will try to get the distributions sorted for a very simplified version of this problem, look up the Bulter-Reed-Dawson algorithm
- WILL NEED to consider a priori that can allow the data to converge so it can be used properly.
- Given book: Principles of Magnetic Resonance Imaging (A Signal Processing Perspective), look at eq 3.69 (the Bloch equation), it is the effect of the applied magnetic field used for MRI
- Project will be done in MATLAB

Action Points For Next Meeting:

- Read the (1053-587x(02)03282-8) IEEE paper, get understanding of what is happening
- Implement Butler-Reed-Dawson algorithm and a simplified version of the method in the paper with 1 dimension for T2