UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITEIT GENT

Bayesian Inference of the 1D Electron Density Profile within the WEST Tokamak using Interferometry

by

Daniel Jordan

A masters thesis presented to the University of Padova
in fulfillment of the thesis requirement for the degree of
Physics of Data

The research was carried out at Ghent University with onsite supervison from
Geert Verdoolaege - Associate Professor
Hao Wu - PhD Student

Remote supervision from the university of Padova was provided by
Lidia Piron - Assistant Professor

2023

Author's Declaration

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

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

Abstract

The abstract should briefly highlight the importance of the research and present its key findings.

Acknowledgements

I would like to thank all the little people who made this thesis possible.

Dedication

This is dedicated to the one I love.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

This document is incomplete. The external file associated with the glossary 'abbreviations' (which should be called `uw-ethesis.gls-abr`) hasn't been created.

This has probably happened because there are no entries defined in this glossary. Did you forget to use `type=abbreviations` when you defined your entries? If you tried to load entries into this glossary with `\loadglsentries` did you remember to use `[abbreviations]` as the optional argument? If you did, check that the definitions in the file you loaded all had the type set to `\glsdefaulttype`.

This message will be removed once the problem has been fixed.

List of Symbols

This document is incomplete. The external file associated with the glossary 'symbols' (which should be called `uw-ethesis.symbols-gls`) hasn't been created.

Check the contents of the file `uw-ethesis.symbols-glo`. If it's empty, that means you haven't indexed any of your entries in this glossary (using commands like `\gls` or `\glsadd`) so this list can't be generated. If the file isn't empty, the document build process hasn't been completed.

You may need to rerun LaTeX. If you already have, it may be that TeX's shell escape doesn't allow you to run makeindex. Check the transcript file `uw-ethesis.log`. If the shell escape is disabled, try one of the following:

- Run the external (Lua) application:

  ```
  makeglossaries-lite "uw-ethesis"
  ```

- Run the external (Perl) application:

  ```
  makeglossaries "uw-ethesis"
  ```

Then rerun LaTeX on this document.

This message will be removed once the problem has been fixed.

Chapter 1

Introduction

- introduce magnetic confinement fusion, the tokamak and the electron density profile
- Justify topic, why the electron density profile is important, why interferometry is a good diagnostic, why Bayesian inference is good in particular GPR. How this can be expanded to combine other diagnostics in an bayesian integrated analysis.

  Quality of confinement ie distinguishes between H and L mode

  Key role in ignition criteria, explain section from Wesson Tokamaks

  Operation limits, Low density electrons can run away, high density over greenwald limit leads to instabilities. I imagine H-mode reduces runaway electrons
- give the details of the WEST tokamak and the IMAS database system.
- explain that NICE exists and that it will be the main source of comparison
- Summarise all work in a few paragraphs
- Give a literature review on what is the current status of the research in this topic.

  Cholinskey Thompson Scattering implementation of GPR.

  Show rough number on papers using GPR in fusion, perhaps a history of it.

  What other diagnostics can be used, how many have papers that use a GPR analysis?

  What are the other inference algorithms for accomplishing the same goal. I imagine for each diagnostic the algorithm is very different whilst bayesian methods

are more widly applicable, GPR can be used for every diagnostic with a linear forward model and monte carlo sampling techniques can be used for all of them, thus simplifying the field.

- Describe what is in the other sections of the thesis

Chapter 2

Background Theory of Bayesian Techniques and WEST Interferometry

This chapter aims to equip the reader with necessary background theory required to reproduce this work and to understand the origin of the inferred electron density profiles presented in the results. It first describes a tokamak fusion device and some relevant physics concepts behind its function. It then describes in a high level manner the inference carried out by Blaise Faugeras and team with their code known as NICE [2].

[2].

After the chapter outlines Bayesian inference and it's connection with GPR. Then the GPR method is explained in detail for a simple regression problem. ~~At this point~~ Interferometry is introduced in enough detail that the GPR method can be altered to infer the electron density profile. The key factor being the new responce matrix. There are various options for advanced alterations that are ~~then~~ explained here and explored in the results section.

Tokamak is a class of fusion device whose name comes from the abbreviation of a Russian phrase which means ~~"~~toroidal chamber with magnetic coils". It consists of a doughnut shaped vacuum chamber surrounded by powerful magnets that aim to confine a high temperature plasma that would otherwise vaporise the chamber. The plasma pressure and temperature are fundamental parameters in the context of nuclear fusion because they dictate the conditions required to overcome the electrostatic repulsion between positively charged atomic nuclei and bring them close enough for the strong nuclear force to initiate

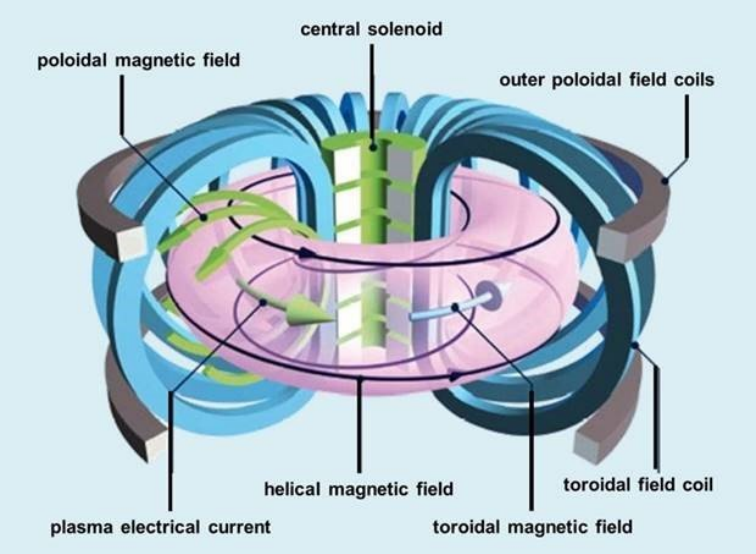

Figure 2.1: ~~Diagram of~~ a tokamak and relevant magnetic fields that create the helical particle trajectory [5].

fusion reactions. In the core of stars like our Sun, the immense pressure and temperature generated by the gravitational collapse create the conditions where hydrogen nuclei (protons) can overcome their natural repulsion and fuse into helium, releasing a tremendous amount of energy in the process. To initiate fusion, hydrogen must be heated to temperatures in the range of tens of millions of degrees Celsius. In a tokamak, this is mainly accomplished with ohmic heating via a driving plasma current and neutral gas injection. This involves accelerating hydrogen ions to high speeds with electric fields and neutralising them the instant before they enter the chamber. The resulting plasma attains the required temperature, allowing nuclei to collide with sufficient energy for fusion reactions to occur. Figure 2.1 shows the position of various magnetic field coils within the tokamak. The toroidal magnetic field exerts an inward force on the plasma thus raising its pressure. A high pressure is required to increase the frequency of collisions so that the energy output could exceed the large heating energy input. The central solenoid induces a current in the plasma which produces the majority of the poloidal magnetic field. This field is essential for confinement but it also plays a key role in plasma stability. The outer poloidal field coils can be controlled in real time to help mitigate instabilities. A real time inference of the electron density profile would assist in identifying instabilities and informing the algorithm that drives the control coils to mitigate them. In addition to high temperature and pressure, the tokamak design seeks to maximize the confinement time of the plasma.

This is essential to allow a sufficient number of fusion reactions to occur before the plasma cools down or loses its stability. The magnetic fields in a tokamak are carefully optimized to prevent rapid plasma loss and minimize heat loss through various mechanisms, including turbulent transport. The shape of the density profile has a large effect on the confinement time.

The combination of the toroidal and poloidal fields shown in figure 2.1 create a helical magnetic field within the plasma. Electrons and ions are accelerated in opposite toroidal directions by the central solenoid yet the both follow a trajectory along the magnetic field lines. This is because a charged particle moving across a magnetic field is succumb to a force perpendicular to its motion. This causes them to gyrate around the magnetic field lines and confines them to follow the magnetic field lines. This is an over simplification although a detailed description of particle motion within a magnetic field is not needed for the purpose of this thesis. It is enough to know that if a single charged particle was within a tokamak then it would perfectly follow a trajectory along the helical path of the magnetic field lines with a small gyration around the field line. When many particles are introduced then collisions can interrupt these trajectories, yet in many models used for data analysis the assumption that particles follow the magnetic field lines is used, including within this thesis.

The magnetic field lines are confined to magnetic flux surfaces, figure 2.2. The toroidal and poloidal flux is constant on magnetic flux surfaces, there is 0 flux across magnetic flux surfaces. Since we assume that the particles follow the magnetic field lines which are strictly bound to these surfaces, we also assume that the density is constant on these surfaces. This allows the density of the entire cross-section to be expressed with a 1D profile as a function of normalised radius ρ, figure 2.3. Where ρ is 0 at the plasma core and 1 at the edge. At the core the poloial magnetic flux is maximum and the edge is the last closed flux surface. Particles past the edge are no longer bound and may interact with the plasma wall. The existance of nested magnetic flux surfaces shown in figure 2.2 rely on the ideal MHD assumptions. Experiments frequently discover magnetic islands which discredits the assumption of nested flux surfaces. The electron density profiles infered rely on this assumption and for many applications, such as real time control a highly accurate infernce is often not required.

NICE is an equilibrium reconstruction code developed for the WEST tokamak. It is relevant because it computes an inference of the electron density profile that is available for comparison to the profile inferred in this work, although NICE's main objective is to infer the shape and position of the magnetic flux surfaces. NICE uses magnetic diagnostics and at WEST these include 421 pickup coils, 36 flux loops and 12 Rogowski coils [7]. The magnetic diagnostics provide the majority of the information. NICE also uses interferom-

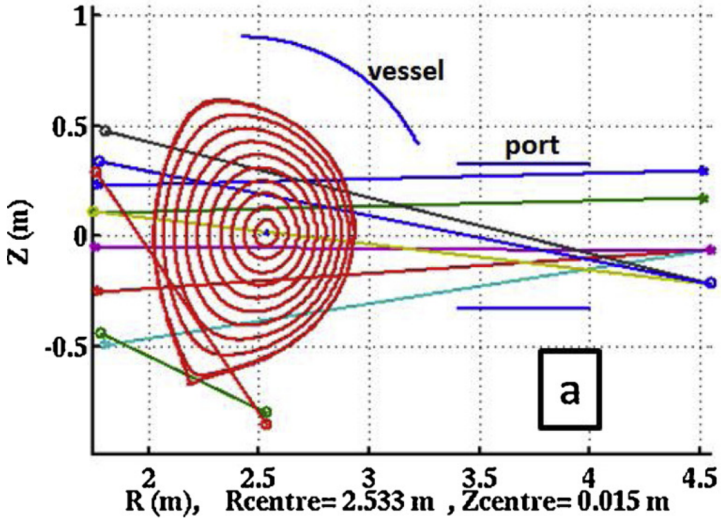

Figure 2.2: The magnetic flux surfaces and interfero-polarimetry laser geometry for a cross-section of the WEST tokamak [3].

etry, polarimetry, motional stark effect and pressure measurements. Section ?? explains how interferometry and polarimetry together can provide information about the poloidal magnetic field, which directly affects the magnetic flux and thus magnetic flux surfaces. NICE performs the inference by minimising a cost function. The cost function determines how well a physical state of the system matches the data received. A state being a specific position and shape of the magnetic flux surfaces and electron density profile. This requires a forward model. The forward model takes a state of the system attempts to compute the signals that would be received by error free diagnostics, if that state was the ground truth. The forward model is a simplified mathematical representation of the measurement process and can never be 100% accurate. This introduces errors into the inference that need to be accounted for. The signals from the forward model can be compared to the actual signals received by the diagnostics in order to compute the cost function. By minimising the cost function the state that best matches the data is found. NICE uses Sequential Quadratic Programming (SQP) as the minimisation algorithm. The optimal state of the system is then stored in the IMAS database for WEST. This includes the 1D electron density profile used as a comparison for the profile inferred in this work ~~using GPR~~. NICE also imposes regularisation terms on their cost function. These penalise the cost function when state properties have features that disagree with prior knowledge. This includes smoothness. We expect the magnetic flux surfaces and electron density profile, to be continuous and

smooth. A state inputted into the cost function that is not smooth triggers the regularisation term which causes the cost function to be larger. Minimising the cost function now also leads to smooth magnetic flux surfaces and electron density profile. This leads into a difficult question, how smooth should it be? They also have a regularisation term to penalise the cost function if the electron density profile is far from 0 at the last closed flux surface or plasma edge. It is prior knowledge that the electron density is near 0 at the plasma edge. How close to 0, and how strong should the regularisation be is still an open question. This work's GPR approach has direct analogs to these regularisation terms, section **??**. Figure 2.3 shows an example of a NICE inferred electron density profile. It is modeled with a cubic spline function. It is the parameters of the cubic spline that are inputted into the cost function. The errors are calculated using a sensitivity method. In short the error is deemed larger for the electron density of a particular normalised radius if a large change in the density leads to a small change in the cost function. In this case we cannot be certain what density is more true because many lead to a similarly low cost function and thus match the data similarly well. To include some more details, the SQP minimisation algorithm computes the hessian of the cost function for minimisation, but this hessian can also be used to measure the sensitivity and thus the errors. The diagonal of the hessian contains the second differential of the cost function for each input parameter. This describes the curvature of the cost function in the direction of each parameter. A smaller curvature means a smaller sensitivity and thus a larger error.

This work aims to use Bayesian inference to obtain the electron density profile. Bayes theorem is written in equation 2.1 for a physical quantity of interest q. The posterior $P(q|D, I)$ is the probability density distribution of q given the measured data D and some prior information I. The maximum of the posterior is the most probable value of q given the data and prior information. The uncertainty of q can also be obtained from the posterior, ~~in this work the posterior will be Gaussian and so the error is simply the standard deviation.~~ The likelihood $P(D|q, I)$ is the probability density function that expresses the probability of the measured data given a fixed value of q and the prior information. The likelihood is described by the experimental error for the data collection. The prior $P(q|I)$ contains information assumed about q before the data is taken. The marginal likelihood or evidence $P(D|I)$ is simply the probability of the data given the prior information only. For posterior computation the marginal likelihood serves as a normalisation factor. Normalisation is often carried out with other means to simplify the posterior computation. Although the marginal likelihood can be used to tune the prior. Sometimes the degree or strength of prior information is uncertain and by finding the strength that maximises the marginal likelihood we find the prior that match's the data the best. Maximising the marginal likelihood to tune ~~the prior~~ also aids to avoid over-fitting ~~the posterior~~ as the trade off

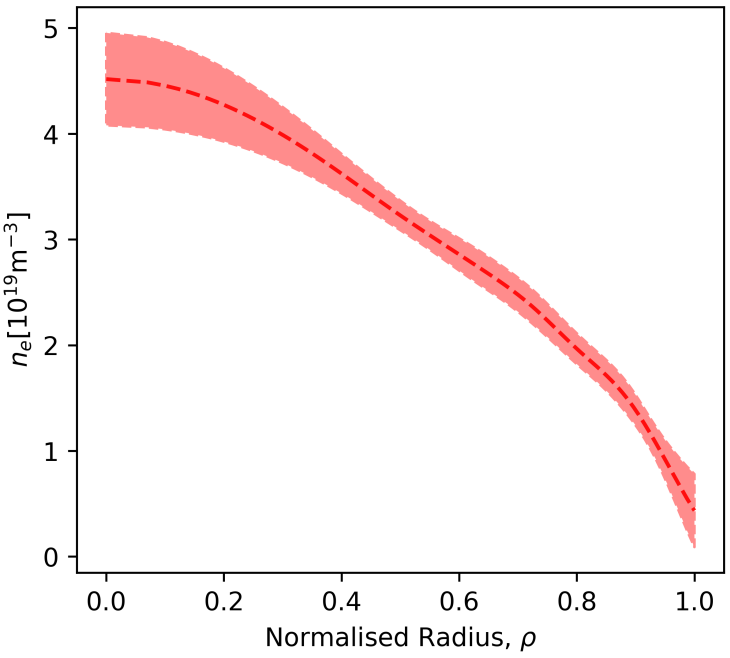

Figure 2.3: Electron density profile inferred by NICE for an instance in time within the WEST tokamak.

between model complexity and data-fit is automatic via the Occam's razor principle [6]. The marginal likelihood method is powerful although it is important to remember that it is not perfect and does not guarantee the most accurate posterior.

$$P(q|D,I) = \frac{P(D|q,I)P(q|I)}{P(D|I)} \tag{2.1}$$

GPR is a form of Bayesian inference where we assume these distributions can be expressed with multivariate Gaussian distributions. When the prior and posterior have the same form then the prior is known as a conjugate prior. This simplifies the inference as it is possible to find a closed form expression of the posterior. Without a closed form expression the posterior must be approximated with sampling techniques. GPR is the technique used in this work to infer the electron density profile from interferometry data. First the technique is introduced for a simple regression problem where we wish to fit a curve to a set of points.

$$\mathcal{N}(\cancel{Y,}\vec{\mu},\overline{\overline{\Sigma}}) = \frac{1}{\sqrt{4\pi^2|\overline{\overline{\Sigma}}|}} e^{-\frac{1}{2}(\vec{y}-\vec{\mu})^T\overline{\overline{\Sigma}}^{-1}(\vec{y}-\vec{\mu})} \tag{2.2}$$

The multivariate Gaussian expression 2.2 can be used to model a curve and its uncertainty, as illustrated in the violin plot 2.4. The mean vector $\vec{\mu}$ holds the y values of the curve at regular intervals along the x axis. The diagonal of the covariance matrix holds the standard deviations of each Gaussian ~~of~~ the multivariate ~~which~~ represents the errors of the curve. Figure 2.4 ~~has~~ 10 Gaussians but in practice many are used to ensure a smooth curve. ~~This will be used to represent the result of GPR~~ which is the most likely curve given the data. ~~The uncertainty of this curve is represented by the standard deviation of the many Gaussians that make it up.~~

The simple regression problem consists of points to which one would like to fit a smooth curve. To solve this using GPR one must define ~~their~~ likelihood ~~(equation 2.5)~~ and prior (equation 2.6) as multivariate Gaussians. The many gaussians that make up these multivariate gaussians are visualised ~~in~~ 2.5. This visualisation method is perfect for the likelihood where the covariance matrix is diagonal. The multiple gaussians that make up the likelihood are not dependant on eachother and so representing them individually provides all the infomation in the likelihood. The prior has a more complex covariance matrix $\overline{\overline{K}}$, and figure 2.5 does not have a complete representation of the priors form. ~~In this simple implimentation the covariance between the prior's Gaussians can be completly encapsulated by a single number known as the lengthscale l in equation 2.6.~~ These visualisations

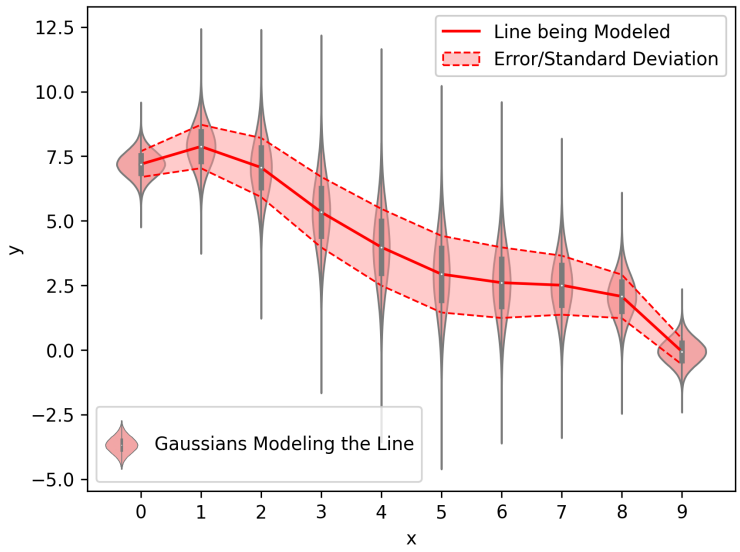

Figure 2.4: Illustrating how many Gaussians can model a line and its uncertainty.

~~of the varius distributions are helpful but~~ to compute the posterior in figure 2.5 one must understand the components of equations 2.3, 2.5, 2.6, and how they can be combined to form the closed form equations 2.4. ~~In the posterior~~ equation 2.3, $\vec{y}$ contains the random variables y_i for regularly spaced positions x_i. Given the data the most probable values of ~~these random variables~~ is expressed in $\vec{\mu}$ and their uncertinites in the diagonal of $\overline{\overline{\Sigma_{post}}}$. In the likelihood equation 2.5 $\vec{D}$, is a vector containing all the ~~data/~~y-values from measurements at various x positions. $\overline{\overline{R}}$ is known as the responce matrix and ensures that the mean vector of the multivariate gaussian $\overline{\overline{R}}\vec{y}$ only contains values of $\vec{(y)}$ at the same x positions the data was recorded at. $\overline{\overline{R}}$ is essentially a filter that removes irrelevant y_i values from $\vec{y}$ for the mean vector of the likelyhood. In the likelihood of figure 2.5 you can see an example of a given $\vec{y}$ and the mean of each gaussian is the value on that line at the same x position as the data point. ϵ is the experimental error of the measurements in $\vec{D}$. In the prior equation 2.6, there are as many gaussians as in the posterior but each of these gaussians have a mean of 0. The $\vec{0}$ symbol represents a vector of zeros the same length as the vector $\vec{y}$ of random variables. To aid visualisation the prior in figure 2.5 only shows 5 out of the 101 gaussians used. The covariance matrix $\overline{\overline{K}}$ is constructed using the kernel K_{ij}. The main role of the amplitude (σ) in the kernel is to set the prior strength. The length scale (l) sets the strength of corelation of neighbouring gaussians. A low length scale ~~that is small relative to the size of the curves features along x~~ means that only gaussians close in

x are highly correlated. Gaussians far have a low correlation meaning they can have a very different mean value. A low length scale allows the fitted curve to have more complexity simililar to a high order polynomial and can lead to overfitting. A high length scale limits the fits ability to curve sharply leading to a simple model similar to a low order polynolial, leading to underfitting. A very high length scale leads to an almost linear fit, every time. This prior is far from perfect. For instance it is often known that the infered values must be positive, for example you cannot have a negative electron density. Since the prior mean vector is set to $\vec{0}$ a negative value is as likely to be infered as a positive value. Since it is gaussian values close to 0 are more likely to be infered than values far from 0. To mitigate this, high amplitude value can be used to lower the prior strength and allow the data in the likelyhod to have more influence on the posterior result. The kernel K_{ij} in equation 2.6 is known as the static exponential square kernel. It is a very commonly used kernel in GPR but far from the only choice. The single value of length scale prevents the inference from haveing long smooth regions with few features followd by regions of high variability. This can be an issue when infering H-mode tokamak plasmas that have a sharp drop off in density at the edge. For these situations a non static kernel can be used that alows the lengthscale to be a function of x which can then allow for posteriors of varying complexity. Regardless of the kernel used, deciding the optimal values of its parameters for a problem is not obvious. A common solution is to use the marginal likelyhood of equation 2.1 as a loss function. The parameters that maxemise it can be found with gradient based methods. The marginal likelihood expresses the probability of the data and so maximising it finds the parameters that maxemises to probability of the data being measured. Although the marginal likelyhood method is also known for automatically deploying Occam's razor principle which finds a balance between closely fitting the data and having a simple model that accounts for the data's errors to have a more accurate inference [6] [8]. Essentially maxemising the marginal likelihood avoids overfitting. Although this method is powerful it does not guarentee to produce parameters that lead to the most accurate fit. To get a more acurate fit, bayesian sampleing techniques can be used, although this is more computationally expensive.

$$\mathcal{N}(\vec{y}, \vec{\mu_{post}}, \overline{\overline{\Sigma_{post}}}) \tag{2.3}$$

$$closedform \tag{2.4}$$

$$\mathcal{N}(\vec{D}, \overline{\overline{R}}\vec{y}, \overline{\overline{\Sigma_{li}}}), \quad \overline{\overline{\Sigma_{li}}} = \vec{\epsilon} I = \begin{vmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_3 \end{vmatrix} \tag{2.5}$$

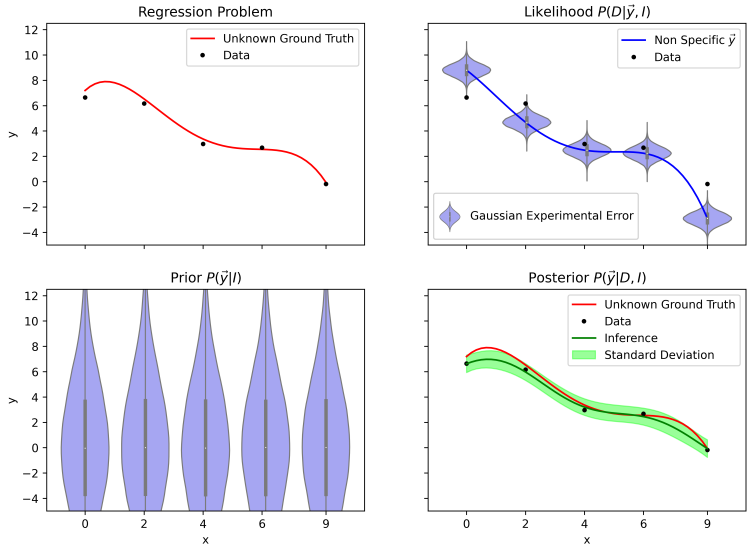

Figure 2.5: A visualisation of the simple GPR process. ~~In truth the prior in this implementation has 101 Gaussians. Only five are shown to aid visualisation. The likelihood truly has the same number of Gaussians as the number of data points. The posterior has 101 Gaussians, their mean is shown by the green line, the standard deviation by the green shade, it is analogous to figure fig:nvg without the Gaussians full form being shown.~~

$$\mathcal{N}(\vec{y}, \vec{0}, \overline{\overline{K}}), \qquad K_{ij} = k(y_i, y_j) = \sigma e^{\frac{y_i - y_j}{2l^2}} \tag{2.6}$$

Interferometry and polarimetry are two diagnostics obtained from the same hardware. Lasers are fired through the plasma at multiple angles. The geometry at WEST is shown in figure 2.2. The refractive index of the plasma is dependant on the electron density. The lasers slow and the phase shift can be measured. This phase shift is proportinal the line integrated electron density n_e, along the line of sight of the lasers, equation 2.7,[4]. The laser wavelength λ, is comnined with other common physical constants to assertain the constant of proportionality. ~~For this reason~~ WEST have stored the line integreated electron density as raw data in the IMAS database. This is the data that will be used for this work.

$$\Delta\phi = \frac{\lambda e^2}{4\pi\epsilon_0 m_e c^2} \int n_e dl \tag{2.7}$$

Polarimetry measures the faraday rotation angle of the lasers. The linearly polarised

lasers experience a rotation as the circularly polarised components travel through the plasma at different speeds. This is due to the small gyration of the electrons around the margnetic field. The faraday rotation angle is proportinal to the line integrated density of $n_e B_{||}$ along the line of sight of the lasers. Where $B_{||}$ is the magnetic field strenghth paralell to the line of sight. Polarimetry has information about the electron density and this work could be extended to bcome a bayesian integrated analysis which includes this infomation in the inference. Currently only interferometry information is used.

$$\theta_F = \frac{\lambda^2 e^3}{8\pi^2 c^3 \epsilon_0 m_e^2} \int n_e B_{||} dl \tag{2.8}$$

To infer the electron density profile with interferometry the GPR process is altered. $\vec{y}$ becomes $\vec{n_e}$ and the likelihood changes. The $\vec{0}$ prior mean can remain the same, the amplitude σ and length-scale l can also be optimised my maximising the marginal likelihood. The data is now in a different space and thus is the likelyhood. The responce matrix R must be created so that it will transform a profile $\vec{n_e}$ into what would be measured by an error free version of the WEST interferometry system given $\vec{n_e}$ is the true profile. The result of $\overline{\overline{R}}\vec{n_e}$ is a vector the same length as the data $\vec{d}$ where each element corresponds to a different interferometry laser or channel.

The responce matrix computation can be summarised in a few steps. NICE provides the magnetic flux at a set of grid points on the tokamak cross section. It also provides the flux at a set of flux surfaces of known normalised radius ρ. A simple 1D interpolation can be used to determine the normalised radius at each grid point. Then using $\vec{n_e}$ another 1D interpolation can be done to determine the electron density at each grid point. After the density at any point along a lasers line of sight can be computed using tringuar mesh interpolation. The density at the golden cross in figure 2.6 can be computed as a weighted sum of the three nearest known density points that form a golden triangle~~, $n_e(t_1), n_e(t_2), n_e(t_3)$.~~

$$n_e(l_i) = \lambda_1 n_e(~~t_1~~) + \lambda_2 n_e(t_2) + \lambda_3 n_e(t_3) \tag{2.9}$$

The line integrated density can be approximated as a sum of many electron densities ~~at N points~~ along the line of sight times the width of their seperation.

$$\int n_e dl \approx \sum_i n_e(l_i)\delta l \tag{2.10}$$

The contribution $w(g_i)$ of each grid point g_i is a sum of all the mesh interpolation coefficiants λ_j used on that point.

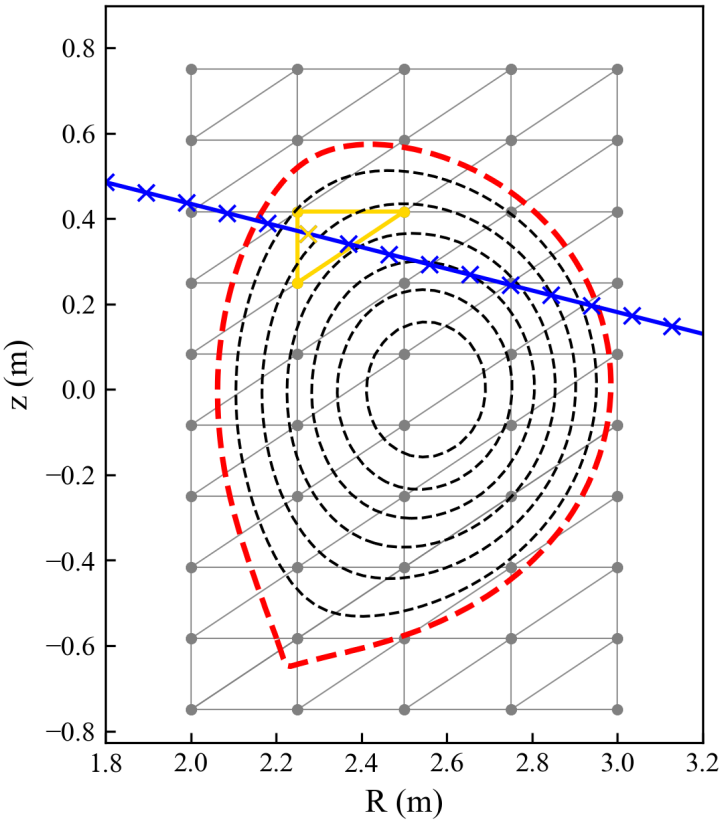

Figure 2.6: An example mesh grid to aid visualisation of the triangular mesh grid interpolation used in the responce matrix construction.

$$\int n_e dl \approx \delta l \sum_i w(g_i) n_e(g_i), w(g_i) = \sum_j \lambda_j \tag{2.11}$$

Each point can be associated to a nearest flux surface f_i ~~equally spaced in ρ~~. This way the contribution $w(f_i)$ of each flux surface is a sum of the contribution at each of its associated grid points g_j.

$$\int n_e dl \approx \delta l \sum_i w(f_i) n_e(f_i), f = \sum_j g_j \tag{2.12}$$

All of these steps equate to a simple rearranging of the origional equation 2.10 to extract the contribution of each flux surface on the final integrated density value. Equation 2.12 can be computed using a vector product.

$$\int n_e dl \approx \delta l \vec{w}^\top \vec{n_e} \tag{2.13}$$

The contribution vector applies to one line of sight. The computation for all lines of sight can be performed by placing the $\vec{w}$ vector for each line of sight as a row in the responce matrix $\overline{\overline{R}}$. Creating a line integrated density vector $\vec{lid}$. The δl is also multiplied into the matrix to simplify notation.

$$\vec{lid} \approx \overline{\overline{R}} \vec{n_e} \tag{2.14}$$

This responce matrix $\overline{\overline{R}}$ can then be used in the closed form expressions **??**, to perfrom a 1D electron density profile inference.

Some further alterations to the GPR method can be made to further increase the reliability of the inference. These include altering the kernel and adding artifical observations to include prior knowledge. The kernel can be changed to a non static kernel. This allows the length scale to change as a function of ρ.

$$K_{ij} = k(\rho_i, \rho_j) = \sigma^2 \left(\frac{2 l(\rho_i) l(\rho_j)}{l(\rho_i)^2 + l(\rho_j)^2} \right)^{1/2} \exp\left(\frac{(\rho_i - \rho_j)^2}{l(\rho_i)^2 + l(\rho_j)^2} \right) \tag{2.15}$$

The length-scale controls smoothness, model complexity and curvature. If these are free to change for different regions of the plasma then there is a greater range of possabilities

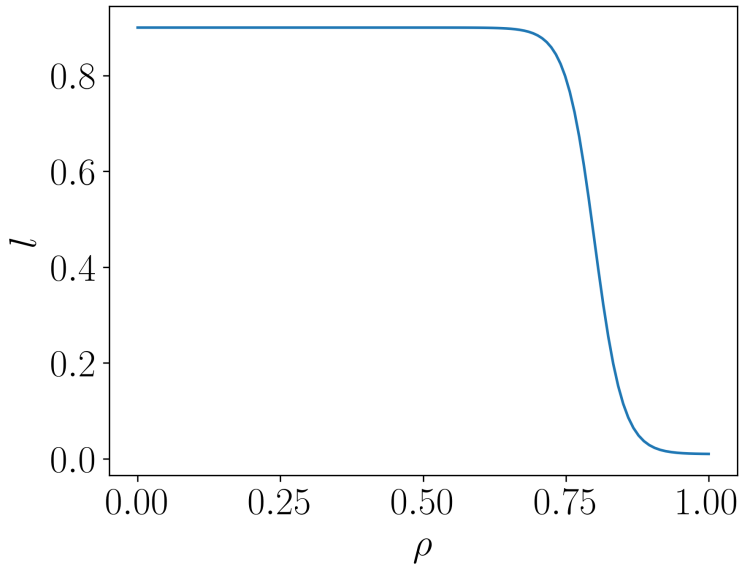

Figure 2.7: Hyperbolic tangent smooth step function for length scale, equation 2.16. Used to capture the drop at the edge of H-mode plasmas [1].

for the final inference. Chilenski used a hypernolic tangent function, equation 2.16, to form a smooth step down from a high length scale at the core to low at the edge [1].

$$l(\rho) = \frac{l_{core} + l_{edge}}{2} + \frac{l_{core} - l_{edge}}{2} tanh\left(\frac{\rho - \rho_{center}}{\rho_{width}}\right) \tag{2.16}$$

The extra freedom at the edge allows the inference to accomodate for a large sudden drop in electron density, which is a common feature for H-mode plasmas. H-mode plasmas are known to have a longer confinement time and thus better fusion performance. WEST does not operate in H-mode, although this method is tested with synthetic data from a simple H-mode simulation.

Put in the cubic spline length scale.

Traditionally prior information should be included in the prior distribution. Although in practice precision errors can make it more convenient to place prior information into the likelyhood in the form of artificial observations. This practise was also adopted by Chilenski [1]. The density is known to be close to 0 at the edge, ($\rho = 1$). It is also known that the density profile is smooth and symetric meaning the gradient of the profile at the core must be close to 0. The point of reflection for the symetry is the core. This

information can be included in the data, $\vec{d}$, with an atrificial experimental error determining the strength of the information included in $\vec{\epsilon}$. Other parts of the GPR method need to be altered to accomodate the new information. The vector to be infered $\vec{a}$ is not only $\vec{n_e}$ but also includes $n_e(\rho = 1)$ and $n'_e(\rho = 0)$ concatonated onto the end. This allows the responce matrix alteration to be simple.

$$\overline{\overline{R}}^{alt} = \begin{vmatrix} \overline{\overline{R}} & \cdots & 0 & 0 \\ \vdots & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{2.17}$$

The prior covariance matrix must also be altered. The covariance between a gradient and non gradient data point is simply the diferntial of the covariance over the gradient data point. For two gradient data points it is a double differential over each point.

$$K'_{ij} = k'(\rho_i, \rho_j) = \frac{\partial k'(\rho'_i, \rho_j)}{\partial \rho'_i} \tag{2.18}$$

$$K''_{ij} = k'(\rho'_i, \rho'_j) = \frac{\partial k''(\rho'_i, \rho'_j)}{\partial \rho'_i \partial \rho'_j} \tag{2.19}$$

In this notation ρ' indicates the position of a gradient data point. The alternate kernal is then.

$$K^{alt} = \begin{matrix} K & K' \\ -K'^{\top} & K'' \end{matrix} \tag{2.20}$$

To summarise the electron density profile is important as it plays a key role in determining the energy confinement time and informing real time control systems. With the assumption of magnetic flux surfaces one can express it as a 1D profile, 2.3. NICE is an equilibrium re-construction code that also infers the electron density profile that can be used as a comparison in this thesis. GPR is a form of bayesian inference that can be applied to interferometry data to infer the electron density profile. A non static kernel can be used to allow the inference to have model complexity that varies with ρ. Prior infomation can be included in the likelihood with artifical observations. In the results section these methods will be deployed on synthetic data. The inference performance can be determined by how closely it fits the ground truth profile. They will also be deployed on real WEST data and the results will be compared the that obtained by NICE.

- Some basic backround theory of a tokamak as a fusion device. Including the schematics of the magnetics and different systems present. Including the magnetic flux surfaces and what they are, assumptions made. What is ρ? Also highlights a little why plasma density profile is important.
- Explain the theory of how NICE did their inference to some degree. ***
- Bayes Theorem for inference problems, each distribution involved and why the margianl likelyhood is useful. ***
- Explian how a multivariate gaussian can model a curve. ***
- Explain Gaussian Process Regression for a generic line fitting problem, including the various distributions constructed and how they are used in the closed form expressions. A little on marginal liklyhood again for parameter optimisation. Mention briefly the alterations to this algorythem that will be made to allow it to be used for interferometry including the responce function change, and virtual observations.
- Explain Interferometry and a little on polarimetry, show west laser geometry
- Explain the magnetic flux surfaces again a little and how this allows us to make the electron density profile 1D again.
- Explain the forward model and how it is computed with the responce matrix.
- Explain how the prior edge and core information can be added and the complications of inserting it into the prior.
- Explain why a non stationary kernel (smooth step) might be beneficial
- summarise the chapter and explain how it ties into the methodology

Chapter 3

Methodology

Explain the details of the how the theory is implimented, what processing power is required, I used python jupyter notebooks numpy, scipy and pytorch. Explain the key parts of the code in order to compute the kernels, marginal likelyhood and inference. What test were done in order to try and improve the inference.

- What format was the data in when it came from west and how did I import it into python
- A bit about the diffent forms the data comes in on the west imas database and which one was chosen.
- How can the accuracy of the forward model be tested and how should this effect the experimental error.
- How does one select an experimental error, is it important for the inference?
- What is the non positive definite matrix error, how can it be avoided, what are the implications of this
- How I used meshgrid to compute the kernel in a fast way
- what are the different methods used for finding the minimum of the marginal likelyhood, scipy minimize, pytorch, grid search. How do they work.
- What distributions did I use to randomly initialize the kernel parameters. Why did I randomly initialise them?

- Why I coded everything with pytorch tensors
- How I optimised the learning rate
- Attempts to use noise
- How did I try and improve the inference.
- summarise the chapter and lead into the results.

Chapter 4

Results

- Make the case as why a 0 mean prior is not always a good Idea, since the marginal likelyhood is not perfect and inferences favour the lower side of NICE, likely due to the 0 prior as when amp is increased the inference will move to NICE but not beyond it. Including prior knowledge when available is always advised.
- Make the case that if I set NICE as the true ground truth and create synthetic data then I get the NICE profile from lowering marginal likelyhood. This shows that everything is set up correctly.
- Show I am getting a local minimum of the marginal likelihood which is nice like and a lower minimum which is parabolic. This could be a result of the marginal likelihood not being a perfect loss function for problems with little data. This seems to be true for SciPy, PyTorch, static and non-static.
- Talk about the discovered more defined ridge, ask for access to log book to see if this is during a H-mode shot. Explain why it might be physically significant but a full monte carlo bayesian should be carried out to verify this.
- Show the static grid search
- Results not yet curated

  Bayesian inference using monte carlo methods

  One size fits all kernel for real time.
- summarise the chapter and lead into conclusions

Chapter 5

Conclusion

The conclusion should have a short summary of each chapter highlighting the main parts of a story from data to inference to insights. Lead into future investigations.

Chapter 6

Future Investigation

- In future the injected gas information could be included to get a prior mean vector. How does this affect the inference? Is the widely accepted 0 mean prior a good idea? Simply make the prior mean high enough so that 99% of the area is in the positive region. Since we know that the electron density is certainly positive.
- Plolarimetry information could be included. Using a thermal profile to get current profile to get polodial magnetic field profile to then use the polarimetry measurements to get more data on the current density. Explain why NICE didn't do this and why this might allow another implementation to improve on NICE.
- try maximizing likelyhood rather than marginal likelyhood for best parameters. See obsidian closely fitting the data.
- Full monte carlo if not done, to get a bayesian ground truth.
- Explain potential for real time inference, why its is important and how you envision it could be done. Using data from many shots and times we could compute the least square error of inference with bayesian ground truth as the loss function to optimise a one size fits all kernel with many parameters. Each l(rho) and sigma(rho) has a different value then the kernel would only have sigma along the diagonal as l can be optimized on the off diagonals to account for different sigma values. This reduces the chance of non positive definite matrix errors. Then using posterior of one inference as prior of next. Why is this method better than a neural network?
- Jeffry finds the parameters are the same for many shots, he thinks I can take the mean and it will work fine as a one size fits all kernel.

- Jeffries Advanced Static Kernel, a reference prior can be used to direct the functional form limitations of a non static kernel which is the Advanced Static Kernel. The reference prior could be NICE, A monte carlo inference for 1 shot, or a parabolic with correct core and edge with edge being 0 and core given by amount of gas injected. It just needs to follow the rough shape as the implimentation is general. Although I fear if I use a parabolic the inferences will be parabolic.

- Jeffries Advanced Static Kernel could be used to help get a one size fits all kernel. You could learn the $sigma_{ij}$ and l_{ij} that lead to the lowest loss over many shots and you don't need to worry about non positive definite errors as $sigma_{ij}$ and l_{ij} are fed into the the Advanced static kernel formula. Then you have very many parameters as

References

[1] M.A. Chilenski, M. Greenwald, Y. Marzouk, N.T. Howard, A.E. White, J.E. Rice, and J.R. Walk. Improved profile fitting and quantification of uncertainty in experimental measurements of impurity transport coefficients using gaussian process regression. *Nuclear fusion*, 55(2):23012–20, 2015.

[2] Blaise Faugeras. An overview of the numerical methods for tokamak plasma equilibrium computation implemented in the nice code. *Fusion engineering and design*, 160:112020–, 2020.

[3] C. Gil, G. Colledani, M. Domenes, D. Volpe, A. Berne, F. Faisse, C. Guillon, J. Morales, P. Moreau, and B. Santraine. Renewal of the interfero-polarimeter diagnostic for west. *Fusion engineering and design*, 140:81–91, 2019.

[4] Ian H. Hutchinson and Ian H. Hutchinson. *Principles of plasma diagnostics / I. H. Hutchinson*. Cambridge University press, Cambridge etc, c1987.

[5] S. Li, H. Jiang, Z. Ren, and C. Xu. Optimal tracking for a divergent-type parabolic pde system in current profile control. *Abstract and applied analysis*, 2014:1–8, 2014.

[6] David J.C. MacKay. *Bayesian methods for adaptive models*. PhD thesis, California Institute of Technology, 1992.

[7] P. Moreau, A. Le-Luyer, P. Spuig, P. Malard, F. Saint-Laurent, J. F. Artaud, J. Morales, B. Faugeras, H. Heumann, B. Cantone, M. Moreau, C. Brun, R. Nouailletas, E. Nardon, B. Santraine, A. Berne, P. Kumari, and S. Belsare. The new magnetic diagnostics in the west tokamak. *Review of scientific instruments*, 89(10):10J109–10J109, 2018.

[8] Carl Edward. Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 2006.

Glossary

This document is incomplete. The external file associated with the glossary 'main' (which should be called `uw-ethesis.gls`) hasn't been created.

Check the contents of the file `uw-ethesis.glo`. If it's empty, that means you haven't indexed any of your entries in this glossary (using commands like `\gls` or `\glsadd`) so this list can't be generated. If the file isn't empty, the document build process hasn't been completed.

You may need to rerun LaTeX. If you already have, it may be that TeX's shell escape doesn't allow you to run makeindex. Check the transcript file `uw-ethesis.log`. If the shell escape is disabled, try one of the following:

- Run the external (Lua) application:

  ```
  makeglossaries-lite "uw-ethesis"
  ```

- Run the external (Perl) application:

  ```
  makeglossaries "uw-ethesis"
  ```

Then rerun LaTeX on this document.

This message will be removed once the problem has been fixed.