University of London
Imperial College London
Department of Plasma Physics

Tokamak Plasma Analysis through Bayesian Diagnostic Modelling

Oliver P. Ford

November 8, 2010

Final University Version
As archived in the University of London library.

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Plasma Physics of the University of London
and the Diploma of Imperial College London

Abstract

This thesis applies the concepts of forward modelling and Bayesian probabilistic analysis to Tokamak plasmas at the Joint European Torus (JET). The techniques are used to greatly improve the accuracy of important plasma parameters and examine plasma physics from existing data that has not been possible with standard methods. Models of the JET interferometry and polarimetry diagnostics are developed and used to asses approximations given in two competing theoretical papers and constitute the first observation of relativistic effects on fusion relevant plasma polarimetry. A highly detailed model is developed for two LIDAR Thomson Scattering diagnostics and combined with the interferometer to infer electron density and temperature profiles with greatly improved accuracy and rigorously quantified uncertainties from all known sources, including uncertainty in the magnetic topology and diagnostic calibrations. The ability to easily cope with a combination of complex systems including unknown calibration parameters is demonstrated. High resolution information inferred about the plasma edge is used to examine the evolution of the H-Mode pedestal in JET plasmas. Bayesian principles are then used to investigate the extent of the information that can be inferred about the plasma current profile from external magnetic diagnostics and the assumption of magneto-hydrodynamic equilibrium. It is shown that a great deal more about the current near the plasma edge can be inferred than was previously thought possible. The presence of the thin layer of high current in the H-mode pedestal is shown and the magnitude variation of the parallel and perpendicular components of this are inferred. The latter is compared to the kinetic measurements, showing surprisingly good agreement.

Contents

List of Figures

0.1. Declaration

I confirm that the work presented in this thesis is entirely my own, except where indicated in the acknowledgements or by references in the main text.

Oliver Ford
28th September 2010

0.2. Acknowledgements

I wish to express my sincere appreciation to my supervisors at JET, Dr Jakob Svensson and Dr Darren McDonald, for giving me the opportunity to work on this project, for supporting me over the past four years and promoting my work amongst the community. Dr Svensson introduced me to both the concepts and practicalities of Bayesian analysis, making the subject of statistics, which had always confused me in the past, suddenly much clearer. I have enjoyed our many long and sometimes deeply philosophical discussions, whether on the topics of inference and plasma physics or the many other besides. Dr McDonald introduced me to much of the world of Tokamaks and Fusion, encouraging me to get involved in the experimental campaigns through which I began to learn a wide range of Tokamak physics, and started to understand the implications and importance of the problems on which I was working.

I would like to thank Dr Alex Meakins for detailed discussions about the LIDAR model but especially for the development, maintainance and continual optimisation of the Genetic Algroithm, in his relentless pursuit of the perfect general optimiser. Without it, many of the investigations would have not been practically possible. Dr Emilia Solano has my gratitude for helping me to understand much of the basic equilibrium theory, encouraging me with the Bayesian equilibrium investigations and helping to interpret the results. Obviously, none of this work would have been possible without the diagnostics, and a good understanding of them, so I would like to thank the diagnostic teams, especially Dr Marc Beurskens, Dr Alexandru Boboc, Dr Joanne Flanagan and Dr Mark Kempenaars for answering all of my queries with precise detail and for providing the data on which this work is based. I would also like to thank Dr Michael Coppins, my contact and supervisor at Imperial College for his advice and direction.

Many others have helped me over my four years at JET and have contributed in small ways to the success of this project: Adrian Capel and Dr David Robson for answering queries and fixing problems with the JET computer systems as I found (or caused) them; Bodhan Beig for discovery of the polarimetry oscillations; my office colleagues David Moulton and Ian Abel and all of the students at Culham for many discussions, and for making life working at Culham more enjoyable; Dr Peter De Vries, Dr Barry Alper, Dr Clive Challis, Dr Joelle Malloux, Dr Elena De La Luna and doubtless many others who I do not recall, for answering all of my questions in the JET control room and helping me to understand the experiments.

Finally I would like to thank my family for their support and for tolerating me frantically trying to meet conference deadlines, while supposedly on holiday.

The work was funded by the EPSRC and by a CASE studentship with UKAEA Culham, was supported by EURATOM and was carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

0.3. Thesis Overview

This thesis applies the concepts of forward modelling and Bayesian probabilistic analysis to Tokamak plasmas at the Joint European Torus (JET), greatly improving the accuracy in the inference of important plasma parameters. The techniques are used to extract important quantities and examine plasma physics from existing data that has not been possible with standard methods. This includes the observation of relativistic effects in plasma polarimetry, high resolution examination of the H-mode pedestal, exploration of experimental equilibria and inference of the pedestal current from magnetic measurements alone.

Chapter 1 introduces fusion, the Tokamak and the basic Tokamak physics relevant to the investigations throughout the rest of the document and chapter 2 introduces the analysis concepts and computation methods and their advantages over standard approaches.

Chapters 3 introduces the specific physics and system details of the Interferometry and Polarimetry diagnostics, then describes the development of their models. Chapter 4 demonstrates the use of the interferometry model to infer electron density profiles with rigorous uncertainties as well as the extraction of magnetic topology information, which is not normally considered, through the assumption of constancy of density on flux surfaces. This is then used with the polarimetry diagnostic model to asses approximations given in two competing theoretical papers. This makes the first observation of relativistic plasma polarimetry effects in fusion relevant plasmas, an effect far smaller than the noise level of the diagnostic and not previously confirmed using standard approaches.

Chapter 5 concerns the two LIDAR Thomson Scattering diagnostics at JET. A highly detailed model is developed including every part of the measurement system and every contribution to the collected data. The two models are combined with the interferometry model and used to infer electron density and temperature profiles with greatly improved accuracy over the standard analysis methods. The ability to easily cope with the combination of complex systems including some uncertain and some unknown calibration parameters is demonstrated, showing that these can be automatically inferred from the data and self-consistency alone. Accurate estimates of other calibration parameters are extracted directly from the data using a verity of methods. A model for the magnetic diagnostics (developed elsewhere) is included in the combination to allow the proper inclusion of magnetic topology uncertainties - a central and recurring problem of Tokamak analysis that is rarely rigorously assessed. High resolution information is inferred about the plasma edge and is used to examine the H-Mode pedestal a vitally important part of Tokamak plasma physics in the progression toward fusion power plants.

In Chapter 6, Bayesian principles are used to investigate what can really be inferred about the plasma current profile from external magnetic diagnostics and the assumption of magneto-hydrodynamic equilibrium - a subject that has been widely debated theoretically. It is shown clearly that a great deal more about the current towards the plasma edge can be inferred than was previously thought possible. The presence of the thin layer of high current in the H-mode pedestal is shown and the magnitude variation of the parallel and perpendicular components of this are inferred. The latter is compared to the kinetic measurements, showing surprisingly good agreement.

0.4. Other Publications of this work

The work in this thesis has been presented at two conferences by myself and is the subject of two published journal articles and one conference proceedings paper. The interferometry and polarimetry modelling of chapters 3 and 4 was presented at the High Temperature Plasma Diagnostics (HTPD) conference, Albuquerque, New Mexico USA in 2008 and was consequently published in Review of Scientific Instruments [1]. The investigation of relativistic plasma polarimetry effects with the JET polarimetry data was published in Plasma Physics And Controlled Fusion (PPCF) [2] in 2009 and this paper was selected for both the Institute of Physics 'IOP Select' and PPCF highlights 2009. The LIDAR Thomson Scattering modelling and initial results were presented at the European Physics Society (EPS) plasma physics conference in Sofia, Bulgaria in 2009 [3].

Dr Jakob Svensson presented some of the results from this work, in an overview of the wider project at the 2010 EPS plasma physics conference [4], on which I was named as a secondary author.

0.5. Notation and Terminology

Throughout this text, vector quantities are denoted by a single underline, e.g. $\underline{\mathrm{x}}$ and matrix quantities with two underlines, e.g. $\underline{\underline{\mathrm{m}}}$. Normal and Gaussian (not normalised) distributions are represented with $\mathcal{G}\left(\underline{\mathrm{x}};\ \underline{\mathrm{x}}_0,\ \underline{\underline{\sigma}}\right)$, denoting that $\underline{\mathrm{x}}$ follows a Gaussian distribution about the mean $\underline{\mathrm{x}}_0$ with covariance matrix given by $\underline{\underline{\sigma}}$. The $\underline{\underline{\sigma}}$ symbol is used for covariance matrices throughout but the scalar σ is used to refer to a standard deviation. For all parts, the raw data of a diagnostic system is represented with D, with a sub/superscript identifying the system.

Chapters 1 and 2 cover the definition of many terms from Bayesian Analysis, from the Tokamak community and some specific to this document. For ease of reference, a glossary is included in Appendix C which gives the meaning and page number of the original definition, for many of these.

1. Introduction

1.1. Fusion

Among the greatest of recent issues is that of energy production. For the past century, this has primarily been supplied by the burning of fossil fuels, but concerns over the effect this has on the climate as well as their rapidly diminishing supply given the predicted growth of world energy requirement has placed considerable emphasis on finding a more sustainable and less damaging alternative.

Over the second half of the last century nuclear fission, which harnesses the energy released from the splitting of a heavy element into two lighter elements, has also been heavily used. While fission has a vastly greater fuel energy density and does not carry the climate change problems of fossil fuels, it has its own set of issues. The raw fuel for nuclear fission is much more abundant than that of fossil fuels but it is still limited, disposal of the highly radioactive waste products of the fission process is extremely difficult and the safety of fission reactors has to be treated carefully. A typical fission reactor will contain enough fuel for a long period of production and failure to control the reaction has in the past resulted in large scale accidents releasing radioactive material into the atmosphere.

Nuclear fusion, where energy is produced from the combination of two light elements into a heavier one, promises to provide a long-term alternative. The process has the same high fuel energy density but, unlike fission, the waste products of the most ideal reactions are not long-term radioactive and the proposed reactor designs contain only a tiny amount of fuel making large scale accidents impossible. The greatest energy release from fusion is achieved by fusing the lightest elements. Given its abundance on earth, hydrogen would be the obvious choice and fusion beginning from hydrogen is the principal source of energy in stars. Unfortunately the reaction involves several stages and the overall cross-section is too low to be of practical terrestrial use. Reactions involving the first two heavy isotopes of hydrogen, deuterium (D) and tritium (T) are more practical, and are shown in table 1.1 along with their output energy. Reactions 1 and 2 occur at equal rates.

$$\begin{array}{lllll}
1) & D+D & \longrightarrow & {}^{3}He \text{ (0.82 MeV)} & +\, n \text{ (2.45 MeV)} \\
2) & D+D & \longrightarrow & T \text{ (1.01 MeV)} & +\, H \text{ (3.02 MeV)} \\
3) & D+T & \longrightarrow & {}^{4}He \text{ (3.50 MeV)} & +\, n \text{ (14.1 MeV)} \\
4) & n+{}^{6}Li & \longrightarrow & T + {}^{4}He & \\
5) & n+{}^{7}Li & \longrightarrow & T + {}^{4}He + n &
\end{array}$$

Table 1.1.: Reactions of use for fusion power generation.

Deuterium is abundant on earth as 1 part in 6500 of hydrogen in typical water. Tritium is hard to find naturally but can be produced from lithium, itself in reasonably plentiful supply, by reactions 4 or 5. It is proposed that this process will take place inside a fusion reactor.

For a mixture of D and/or T in thermal equilibrium, reactions can be initiated by the thermal motion of the particles. This is known as thermonuclear fusion and the total output power is given by:

$$P_{fusion} = (\text{Reaction rate}) \times E = n_1 n_2 \left\langle \sigma v \right\rangle E \tag{1.1}$$

Where the n_i are the density of reacting species, $\left\langle \sigma v \right\rangle$ is the thermal average of the product of velocity and cross-section and E is the usable energy released from each reaction.

To obtain a useful power output, thermonuclear fusion requires either very high density (as is found in the core of the Sun), or very high average cross-section and hence high temperature. For a self sustaining reaction the generated power needs to be at least equal to the power being lost from the fuel mixture. This is called the point of 'ignition'. Typically, the power loss is dependent on the total thermal energy W of the fuel mixture and this dependence is characterised by the energy confinement time τ_E by $P_{loss} = \frac{W}{\tau_E}$. Balancing these leads to equation 1.2

$$n_1 n_2 \left\langle \sigma v \right\rangle E \geq \frac{3\left(n_e + n_i\right) k_B T}{2\tau_E} \tag{1.2}$$

For a 50-50 mix of D or T, $n_e = n_i = 2n_1 = 2n_2$ and so this becomes:

$$n_e \tau_E \geq \frac{12}{E} \frac{k_B T}{\left\langle \sigma v \right\rangle} \tag{1.3}$$

This requirement on density and confinement time for any given temperature was first derived by Lawson [5] and is widely known as the Lawson Criterion. The temperature which gives the lowest requirement is where $k_B T / \left\langle \sigma v \right\rangle$ is a minimum. This is $2 \times 10^8 K$ for D–D and $3 \times 10^7 K$ for D–T. The term E is simply the energy output of a single reaction but unfortunately, while the energy released in the neutral products can be harnessed, it is difficult to keep them inside the reacting region of any reactor and so it can not contribute to heating the fuel. This means E must be set only to the charged product energies. With this included, equation 1.4 gives Lawson's criterion for D-T and D-D at these temperatures.

$$\begin{aligned} \text{D-D:} \quad & n_e \tau_E \geq 10^{22} s m^{-3} [5] \\ \text{D-T (50/50 mix):} \quad & n_e \tau_E \geq 10^{20} s m^{-3} \end{aligned} \tag{1.4}$$

It is clear the D-T would be the preferable choice due to the lower optimum temperature and lower requirement at this temperature. However, a complication arises when τ_E depends on

T which is often the case in proposed reactors. It is then convenient to find a requirement on n and $\tau_E\left(T\right)$ for ignition which does not itself depend on T. The thermonuclear cross-section for $D{-}T$ can be approximated by $\langle\sigma v\rangle \approx 1.1\times10^{-24}T^2$ [6] with T expressed in keV (and k_B included in the constant). This approximation is valid between 10keV and 20keV. Entering this into equation 1.3 gives, for $D-T$:

$$n_e\tau_E T \geq 3\times10^{21}m^{-3}skeV \tag{1.5}$$

The left hand side is known as the fusion triple product and the triple product achieved by an experimental reactor is often used to asses how close to fusion ignition it is.

1.2. The Tokamak

The temperatures required for fusion are so high that contact with any solid material surfaces used to contain the reacting fuels would cause intolerable damage to those surfaces. There are three ways in which the reacting fuels can feasibly be contained without solid materials. Gravitation confinement occurs in stars, where the fuels are naturally compressed to high density and temperature by the star's own gravity. In inertial confinement, high power lasers are used to rapidly heat a small capsule of solid fuel, either directly or via x-rays generated by heating an enclosure. Fusion takes place in the high pressure created before the fuel has time to expand. Finally, in magnetic confinement fusion, gaseous fuels are ionised to become a plasma and the charged particles confined using magnetic fields.

Initially, cylindrical 'Z-pinches' were developed in which the plasma is forced toward the centre by a plasma current driven along the cylinder. To avoid losses at the ends, the cylinder is bent around to meet itself and form a toroidal pinch in which the pinch effect is provided by a current driven around the torus. A metal loop passes through the centre of the torus and through an external winding in which a current is driven, driving the *plasma current* I_p in a similar way to a transformer. The plasma current creates a poloidal magnetic field B_θ as shown in figure 1.1a and the movement of the charged particles in this field creates a force toward the centre of the poloidal cross-section of the plasma. This configuration is very unstable but can be stabilised by adding a field in the toroidal direction B_t using external coils as shown in 1.1b. The combination of the two fields is the basic principle of the tokamak.

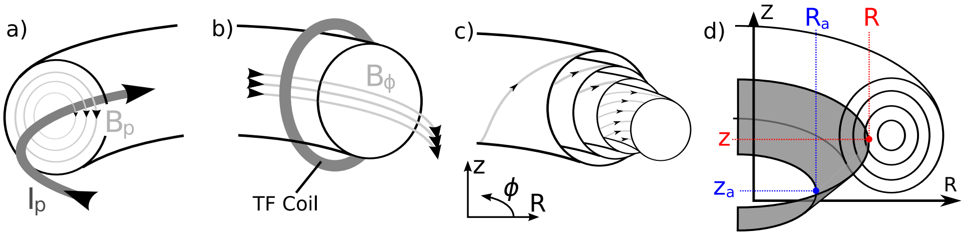

Figure 1.1.: Basic magnetic configuration of the Tokamak. a) Poloidal field created by plasma current. b) Toroidal field created by external coils. c) Combination of the two resulting in nested magnetic surfaces. d) Definition of poloidal magnetic flux $\psi(R, Z)$

For most tokamaks, there are many *toroidal field (TF) coils* and the variation in the toroidal field around the tokamak, known as *ripple*, is very small (less than 1%). The field is usually approximated as a toroidally symmetric average field B_ϕ. The combination of this and B_θ results in toroidally symmetric helical field lines which pass around the torus eventually joining up on themselves. These lines describe an infinite set of nested toroidal surfaces as shown in

1.1c, along with the coordinate system (R, ϕ, Z) usually used in experimental tokamak physics.

The B_ϕ field can also have a contribution from poloidal currents inside the plasma, which increase or decrease that from the TF coils, which is known as the *vacuum toroidal field.*

Flux Surfaces

It is useful to define the *poloidal magnetic flux* $\psi(R, Z)$, as the integral of the poloidal field through a surface (shown in figure 1.1d) from some fixed reference point (R_a, Z_a) in the poloidal plane to the point in question (R, Z), covering 1 radian in the toroidal direction. For the toroidally symmetric tokamak, ψ is constant everywhere on each magnetic surface, giving the surfaces the name *flux surfaces.* The choice of (R_a, Z_a) changes ψ only by a constant value and is usually chosen as $(0, 0)$. The outermost complete flux surface which does not come into contact with a physical surface is known as the *last closed flux surface (LCFS)* and the centre of the set is known as the *magnetic axis.* The average pitch of the magnetic field on each surface is related to the *safety factor* q which can be defined as the number of times a field line on the surface goes around toroidally for a given number of times poloidally. The name is given because the value has important effects on the stability of the configuration. Surfaces where q is a low order rational fraction play a special role as individual field lines cover only part of the surface.

Shaping and configuration

The poloidal field also has a large contribution from several external shaping and stabilisation coils, known as *PF* coils, which also carry a toroidal current. These are used to control the position and shape of the main bulk of the plasma, usually approximately described in terms of its position (R_0, Z_0), minor radius a, *elongation (κ)* and triangularity (δ). The PF coils are also used to react to oppose a violent instability where the plasma undergoes a rapid vertical displacement.

In the simplest configuration, shown in figure 1.2a, the plasma boundary is defined by direct contact with the solid material walls. This will be somewhere on the *first wall*, which is the path in the poloidal plane which traces the nearest material surface, at any toroidal location, to the approximately axisymmetric plasma. The majority of plasma leaving the bulk will impact this part of the wall, causing significant heat, damage and sputtering where neutral particles of the solid material are released and migrate into the bulk plasma. These *impurities* are then ionised by collisions with the plasma. The high Z elements become only partially ionised and radiate a large part of the plasma energy, cooling it significantly. To reduce this, the configuration is set so that first contact is made with a special set of tiles made from materials either resilient to sputtering or of low-Z elements - called a *limiter*.

Many tokamaks have a set of PF coils close to the top and/or bottom of the vessel which can be used to create a zero point in the total poloidal field (also known as an *X-point*). This is used to create a diverted plasma, where the LCFS passes through the X-point and is known as the *separatrix*. In this configuration, the first open surface contacts the first wall away from the bulk of the plasma. This reduces the impurities entering the plasma. In many modern tokamaks, these *strike points* are arranged to be in a partially enclosed part of the vessel, called the *divertor*, which further reduces impurities. Figure 1.2b and c show the flux surfaces for the diverted configuration and the names of the parts of the plasma including the surfaces just outside the LCFS through which charged plasma leaves the bulk plasma, known as the *scrape off layer (SOL)*.

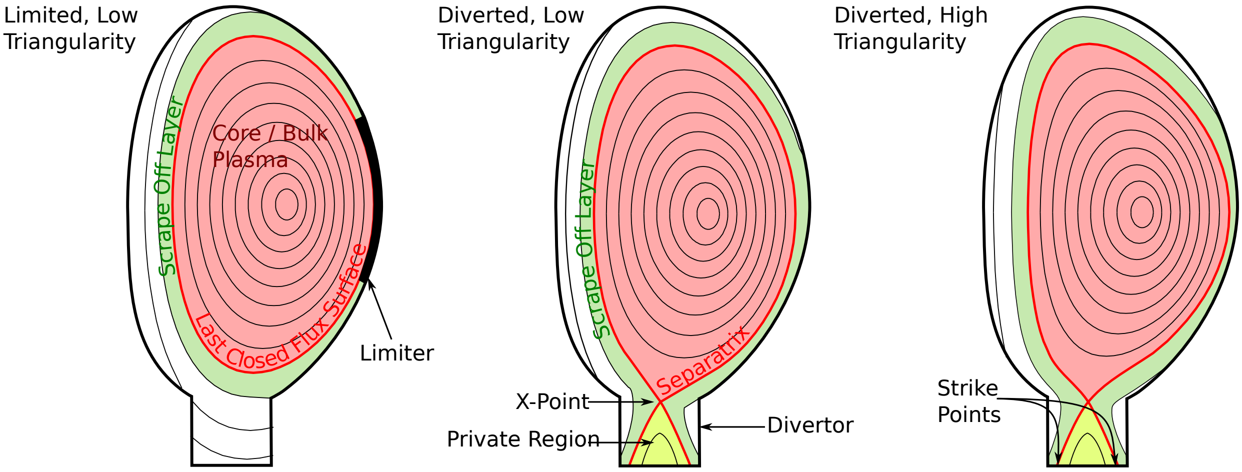

Figure 1.2.: a) Limited plasma, b) Low triangularity divertor plasma, c) High triangularity divertor plasma.

1.3. Tokamak Equilibrium

1.3.1. The poloidal current flux - f

As described in the previous section, the poloidal magnetic flux $\psi(R, Z)$ is defined in the terms of the integral of the poloidal field $\underline{\mathrm{B}_\theta} = (B_R, 0, B_Z)$ passing through a toroidally symmetric surface, defined by a line from a fixed point to (R, Z) in the poloidal plane. The poloidal field can therefore be written as:

$$\underline{\mathrm{B}_\theta} = \frac{1}{R}\nabla\psi \times \underline{\hat{\phi}} \tag{1.6}$$

The poloidal current flux $f(R, Z)$ can be defined in the same way and the poloidal current $\underline{\mathrm{J}_\theta}$ written as:

$$\begin{aligned} \underline{\mathrm{J}_\theta} &= \frac{1}{R}\nabla f \times \underline{\hat{\phi}} && (1.7)\\ &= -\frac{1}{R}\frac{\partial f}{\partial Z}\underline{\hat{\mathrm{R}}} + \frac{1}{R}\frac{\partial f}{\partial R}\underline{\hat{\mathrm{Z}}} && (1.8) \end{aligned}$$

For slowly time-varying plasmas, where there are no large rapidly changing electric fields, $\underline{\mathrm{B}}$ and $\underline{\mathrm{J}}$ are related via Ampere's law. For toroidally symmetric plasmas ($\partial/\partial\phi = 0$), this is:

$$\mu_0\underline{\mathrm{J}} = \nabla \times \underline{\mathrm{B}} = -\frac{\partial B_\phi}{\partial Z}\underline{\hat{\mathrm{R}}} + \left(\frac{\partial B_R}{\partial Z} - \frac{\partial B_Z}{\partial R}\right)\underline{\hat{\phi}} + \frac{1}{R}\frac{\partial\left(RB_\phi\right)}{\partial R}\underline{\hat{\mathrm{Z}}} \tag{1.9}$$

Equating the poloidal components of this with $\underline{\mathrm{J}_\theta}$ from equation 1.8 gives the relation between f and the toroidal field:

$$\mu_0 f = RB_\phi \tag{1.10}$$

With no plasma, the current flux is constant everywhere inside the toroidal field coils $f = f_0$ giving the vacuum toroidal field $B_\phi^{\mathrm{vac}} = \mu_0 f_0/R$.

1.3.2. Low flow MHD equilibrium

Under the assumptions of Magneto-hydrodynamics (MHD) and assuming an isotropic scalar pressure p, the evolution of the plasma conditions is given by the MHD momentum equation:

$$\rho\left[\frac{\partial\underline{\mathrm{v}}}{\partial t} + \left(\underline{\mathrm{v}}\cdot\nabla\right)\underline{\mathrm{v}}\right] = \underline{\mathrm{J}} \times \underline{\mathrm{B}} - \nabla p \tag{1.11}$$

One of the simplest but most useful observations of the plasma in machines like JET is that the plasma is usually stable and the plasma density, temperature and most other quantities do not experience large rapid changes. The time derivatives $\partial/\partial t$ must therefore be small. It is also common to assume that the plasma fluid velocity $\underline{v}$ is small, in which case the momentum equation reduces to implying a balance between the Lorentz force and the pressure gradient force:

$$\underline{J} \times \underline{B} \simeq \nabla p \tag{1.12}$$

A derivation including both toroidal and poloidal flow can be found elsewhere[7], from which it can be seen that the assumption of no flow is valid if the toroidal and poloidal velocities are subsonic. For the JET tokamak, which is studied in this work, toroidal velocities observed are $v_\phi \sim 100km/s$ with Mach numbers (ratio of fluid velocity to thermal velocity) of at most ~ 0.5 [8] and the poloidal flows are much smaller (usually $v_\theta < 10km/s$[9]).

Equation 1.12 implies there can be no variation in pressure along field lines since $\underline{B} \cdot \nabla p = 0$. For most flux surfaces the field line visits the entire surface implying that the pressure must be constant on each surface. For plasmas with little flow, pressure can therefore be written as a function of the magnetic flux $p(\psi)$ or more commonly, of the *normalised flux* ψ_N, which is normalised to 0 at the magnetic axis and to 1 at the LCFS. Similarly $\underline{J} \cdot \nabla p = 0$ implies that there is no current perpendicular to the surfaces and so f may be written $f(\psi_N)$. In almost all real plasmas, the magnetic flux ψ follows a monotonic function from the magnetic axis to the LCFS making both $p(\psi_N)$ and $f(\psi_N)$ single-valued functions, though not necessarily monotonic themselves.

Splitting equation 1.12 into toroidal and poloidal components and inserting the relations for J_θ to f and B_θ to ψ gives:

$$\left(\frac{1}{R}\nabla f \ \times \ \hat{\underline{\phi}}\right) \times B_\phi \hat{\underline{\phi}} + J_\phi \hat{\underline{\phi}} \times \left(\frac{1}{R}\nabla \psi \ \times \ \hat{\underline{\phi}}\right) \simeq \nabla p \tag{1.13}$$

Rewriting the derivatives of p and f in terms of the magnetic flux $\nabla f = \frac{\partial f}{\partial \psi}\nabla\psi$ and $\nabla p = \frac{\partial p}{\partial \psi}\nabla\psi$ and inserting these and equation 1.10 into equation 1.13 gives the force balance in terms of only the toroidal current J_ϕ and the flux functions f and p.

$$\begin{aligned} -\frac{1}{R}\frac{\partial f}{\partial \psi}\frac{\mu_0 f}{R}\nabla\psi + \frac{1}{R}J_\phi \nabla\psi &\simeq \frac{\partial p}{\partial \psi}\nabla\psi \\ \frac{\nabla\psi}{R}\left(J_\phi - \frac{\mu_0}{R}ff' - Rp'\right) &\simeq 0 \end{aligned} \tag{1.14}$$

If the equality is exact and the solution non-trivial, the $\nabla\psi/R$ terms are dropped, giving:

$$J_\phi = Rp' + \frac{\mu_0}{R}ff' \tag{1.15}$$

In this work, the toroidal current J_ϕ is treated as an underlying physical parameter, as it relates directly to the diagnostics being analysed and ψ is calculated as a linear function of J (detailed in section 2.5). More commonly in tokamak physics, the toroidal component of Ampere's law (equation 1.9) is used to rewrite J_ϕ as a derivative of ψ which gives the well known Grad-Shafranov[10, 11, 12] (GS) equation for flow-free, isotropic pressure equilibria:

$$-\frac{1}{R}\frac{\partial^2\psi}{\partial Z^2} - \frac{\partial}{\partial R}\left(\frac{1}{R}\frac{\partial\psi}{\partial R}\right) = \mu_0 R\frac{\partial p}{\partial\psi} + \frac{\mu_0^2}{R} f\frac{\partial f}{\partial\psi} \tag{1.16}$$

1.3.3. Grad-Shafranov solutions and EFIT

It is normal in tokamak analysis to attempt to solve the GS equation for ψ subject to boundary conditions given by measures of ψ and $\nabla\psi$ from magnetic pick-ups and loops outside the plasma. This requires some prescription of $p'(\psi)$ and $ff'(\psi)$ since there is in general no unique solution if ψ, p' and ff' are all completely unknown. In many cases, if these are assigned functions of only a few parameters, a solution to them and ψ can be found[13]. It is not entirely clear in what situations such solutions are unique or what effects the crude ad-hoc assumptions made about the profile shape have on the solution. The solution will not in general match the boundary conditions measurements exactly and the disagreement is assumed to be entirely due to measurement inaccuracy.

One of the most common codes used to obtain GS solutions is known as *EFIT* (Equilibrium FITing) which employs a two stage iterative procedure. In the first stage, the parameters of $p'(\psi_0)$ and $ff'(\psi_0)$ based on fixed $\psi_0(R, Z)$ are determined by a linear fit of J_ϕ given by equation 1.15 to the magnetic data. The second stage then solves equation 1.16 with these fixed p' and ff' profiles. Functional forms based on low order polynomials or a spline with low number of knots are available and the solution can also be constrained by internal magnetic or pressure information from other diagnostics or a-priori constraints like the specification of the central safety factor $q(0)$. At JET, a special version EFITJ[14] is employed which includes the effects of the ferromagnetic *iron core* of the JET transformer structure.

1.3.4. Current and pressure moments

Despite the ill-defined nature of the equilibrium problem, some quantities derived from GS solutions are always accurate, especially those that relate to low order moments of pressure and current inside the boundary. These volume/area integrals relate directly to the line integral of the poloidal magnetic field or flux around the boundary which is well determined by the external magnetic measurements. The following are moments of the toroidal current density:

- Plasma Current: $I_p = \int J_\phi \, dA$

  The toroidal current density of the plasma integrated over the whole poloidal plane.

- Current centroid position $R_{J0} = \frac{1}{I_p} \int R J_\phi \, dA, \qquad Z_{J0} = \frac{1}{I_p} \int Z J_\phi \, dA$

  The current density weighted average position in the poloidal plane.

There are also three well determined quantities that also involve pressure and poloidal current, known as the *Shafranov integrals*[15, 13] s_1, s_2 and s_3. These provide constraints which relate to the following physically relevant quantities:

- Average *Poloidal Beta* $\beta_\theta = \frac{4}{\mu_0 r_c I_p^2} \int p dV$

  The ratio of the volume integrated plasma pressure to poloidal magnetic field energy.

- Magnetisation $\mu = \int (B_\phi - B_{\mathrm{vac}})/2\mu_0 \, dV$

  Related to energy stored in the increase or decrease of vacuum toroidal field by the plasma poloidal current.

- Magnetic inductance $l_I = \frac{1}{I_p^2} \int B_\theta^2/2\mu_0 \, dV$

  Related to the energy stored in the poloidal field created by the toroidal plasma current.

The first Shafranov integral s_1 relates approximately to $3\beta_\theta - \mu + l_I$, the second s_2 to $\beta_\theta + \mu + l_I$ and the third s_3 to $2\beta_\theta + l_I$. It is clear that $\beta_\theta + l_I/2$ can be determined from these but not β_θ or l_i independently. There has been much discussion of what can be known about the current and pressure inside the plasma boundary beyond these quantities. Analytical calculations [16, 17] are restricted to simple cases and show, regardless of how accurate the magnetic measurements are, that only I_p, s_1 and s_2 can be determined in an entirely cylindrical plasma and that only a single higher moment can be known in the circular cross-section toroidal case. These proofs do not hold for non-circular plasmas but it is argued qualitatively that higher order moments will effect the magnetic measurements by too little to be practically measured. Such arguments do not quantitatively asses any real uncertainties and usually overlook the accuracy in the higher order moments that comes from having so many magnetic sensors (hundreds in the cases examined here) all measuring the same plasma with independent random error. More recently, the effect of the real uncertainties has begun to be investigated, such as by assessing linear perturbations to p' and ff', once an equilibrium solution has been found[18]. Experimentally and numerically, it has been shown [19] that for highly elongated (or possibly otherwise strongly shaped) plasmas that, at the very least, β_θ and l_i can be identified separately from the external poloidal magnetic field. In other cases, it is possible to make an independent measurement (e.g. μ taken from a diamagnetic loop) to separate these two moments.

An important effect of high β_θ is that it causes a shift towards larger R of the magnetic axis relative to the centre of the LCFS. This is known as the *Shafranov shift* and is approximately proportional to β_θ. An equilibrium inference with significantly incorrect β_θ will present a significantly incorrect flux surface topology in the plasma core. If, as if often the case, these surfaces are used as a coordinate system upon which to infer other quantities by assuming those quantities are constant on flux surface, they can introduce large systematic errors and apparent disagreement between different diagnostics.

1.3.5. Other flux functions

In between collisions, the plasma particles are approximately confined to gyrate around field lines but are free to travel along them. The low density and high temperatures in tokamak plasmas mean that the particles travel far along field lines before colliding and will explore most of a flux surface between collisions. Any local input of heat will rapidly spread around the entire surface and so, on all but very short time-scales, the temperature can be assumed constant on any flux surface. With the earlier assumption that pressure is also constant, it follows that density can be assumed to be a flux function.

1.4. The High Confinement Mode (H-Mode)

1.4.1. Transport and Confinement

While the variation of particle density and temperature has been assumed small within each surface, equilibrium implies nothing about the variation between neighbouring flux surfaces as arbitrarily large pressure gradients can be supported by sufficient local current. In reality, a small flow of energy and particles between surfaces is also present which is known generally as *cross-field transport.* The two simplest models for this are *collisional transport*, which includes the transfer caused by electron-ion collisions and *neoclassical transport*, which also includes the effects of particles reflected by the increasing field strength on the inboard side of the tokamak and trapped in banana-shaped orbits. Experimentally, much greater cross-field transport rates are observed which, due to the unknown mechanism, is termed *anomalous transport.* Theoretical and experimental studies indicate that it is due to turbulent transport, where small scale fluctuations/eddies of the equilibrium quantities transfer particles and energy across the average/equilibrium surfaces.

In almost every case, the transport rates increase with increasing density/temperature gradient and act to reduce the gradient, resulting in smooth equilibrium profiles. The plasma heating is almost always concentrated on the plasma core which gives temperature profiles that typically fall smoothly and monotonically towards to edge. Density profiles, however, are not necessarily simple or monotonic. The different local transport rates through a given plasma combine to give the plasma's overall particle and energy confinement which relates directly to the fusion yield and power production as described in section 1.1.

1.4.2. Confinement Modes

Early in tokamak research, it was discovered that when increasing input heating power past a critical level, the global energy confinement improved abruptly[20]. Operation above and below this threshold became known as the *Low* and *High* confinement modes (L-Mode and H-Mode). The higher confinement in H-Mode was found to be mainly due a region of abnormally low transport close to the plasma edge known as the *edge transport barrier* (ETB)[21]. Many experiments have been conducted to obtain empirical expressions for the plasma conditions required to enter H-Mode and many theories proposed to explain the barrier. At the time of writing, the most widely accepted proposes the suppression of turbulent transport by a large shear in the plasma flow in the barrier region. However, the cause and mechanisms are not well understood and the barrier characteristics (extent, size, gradients etc) cannot be predicted at present.

1.4.3. Effects of the Edge Transport Barrier

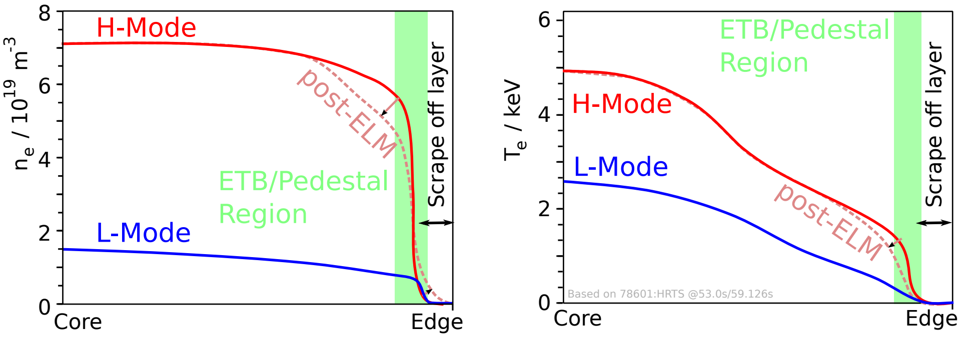

Figure 1.3.: Typical electron density and temperature profiles in low confinement mode (blue) and in high confinement mode (red) before (solid) and after (dashed) an edge localised mode crash.

Figure 1.3 shows typical tokamak electron density n_e and temperature T_e profiles in L-Mode and in H-Mode. The steep gradient of the ETB can be seen clearly in the H-Mode T_e profile and while the gradient remains roughly the same as for the L-Mode in the core, the whole profile is raised on the *pedestal* of the edge region. The n_e profile shows a much higher and steeper pedestal in H-Mode but the pedestal feature is often observed in L-mode, possibly due to the higher charged particle source (due to high neutral density) in the edge region.

The ETB/pedestal region exhibits distinct features in many plasma quantities and to measure these the diagnostics are often designed to have high spatial resolution at the plasma edge. Data analysis is also effected as smoothing, regularisation and fitting techniques tend to underestimate the pedestal gradient unless great care is taken. The large n_e and T_e gradients imply a large pressure gradient which, through the equilibrium, suggests a narrow field-perpendicular current density in the ETB region that is much larger than in the plasma core. Models for the parallel current also give a positive dependence on n_e and T_e gradients and so predict a large parallel current density. The equilibrium fitting procedure described in section 1.3 usually has insufficient freedom in the parametrisation to correctly describe the pressure pedestal and edge current.

Reduced transport in the ETB clearly raises the plasma temperature and density across the entire plasma volume but it has also been shown experimentally [22, 23] that reduced transport in one region can also reduce it neighbouring regions, giving the profile an apparent rigidity of shape known as *stiffness* and further improving the confinement. For many H-mode plasmas, a large part of the overall confinement can be due to the presence of the ETB, making it of great importance to the final goal of tokamak research.

1.4.4. Edge Localised Modes (ELMs)

Another feature of the H-Mode is the *edge localised mode* (ELM), a type of instability which results in a periodic (possibly only partial) collapse of the pedestal as shown dashed in figure 1.3. A large amount of plasma is rapidly expelled into the SOL and can be observed by the recombination emission as it is neutralised on or near the first wall. After the collapse, the pedestal gradually rebuilds until the next ELM occurs. This inter-ELM period, which is usually fairly regular for a given set of the global plasma parameters, is used to distinguish between two main types of ELM. With increasing heating power, *Type-I* ELMs increase in frequency and *Type-III* ELMs decrease in frequency [24]. Figure 1.4 shows the D-alpha emission intensity from near the SOL/first-wall contact point for a typical Type-I ELMy H-mode and a typical Type-III ELMy H-Mode. The Type-I ELMs are larger and lower frequency than Type-III ELMs which are often called *grassy* ELMs due to their appearance on the emission signal.

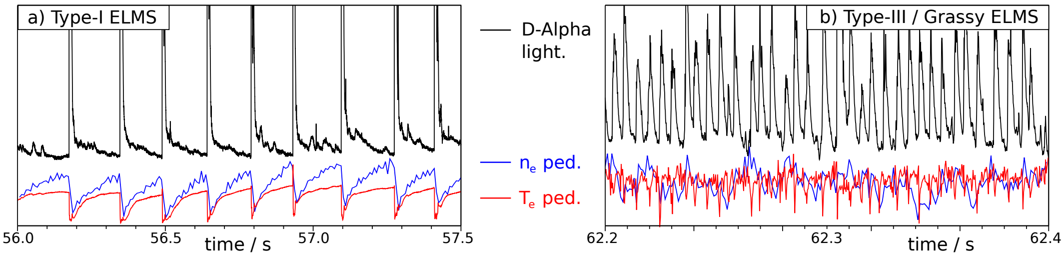

Figure 1.4.: Evolution of the deuterium-alpha line emission intensity (black) and electron density (blue) and temperature (red) at the top of the pedestal during a) Type-I and b) Type-III ELM crashes.

The cause and mechanism for ELMs is not entirely understood. The current leading theory attributes at least Type-I ELMs to either a current instability called a Peeling mode, a pressure driven instability called a Ballooning mode or possibly a combination of the two. Testing of the Peeling/Ballooning model is a highly active area of research as the energy expelled during a Type-I ELM crash would cause significant damage to the first wall in reactor conditions. It will be essential to be able to predict when an ELM crash will occur, determine how much energy will be ejected and possibly suppress or prematurely trigger ELMs. Because the pedestal is such a spatially localised phenomena and the ELM period so short, even measuring the pedestal quantities between ELMs is difficult. Despite the importance of the current in the Peeling/Ballooning model, no reliable measurements have yet been made of the evolution of the pedestal current within the ELM cycle and in most analyses, the pedestal current is imposed based on theoretical predictions.

1.5. The Joint European Torus (JET)

The *Joint European Torus* (JET), from which the data this work is based on comes, is a large Tokamak with a major radius of 2.96m and a minor radius of around 0.9m (horizontal). It was built in the early 1980s and was designed to investigate plasmas approaching fusion reactor relevant plasma conditions. At the time of writing, it is the largest Tokamak ever built and holds the record for the highest fusion power of 16MW, a ratio to input heating power of $Q \approx 0.7$. This was achieved using a 50/50 mix of deuterium and tritium and JET is currently the only tokamak capable of handling full operation with tritium fuel, although it usually operates with D-D plasmas to limit activation of the vessel and components. All of the plasmas analysed in this work are principally Deuterium. Present day investigations at JET play a major role in the design of *ITER* ('The Way' in Latin and originally the 'International Thermonuclear Experimental Reactor'), a new Tokamak that will provide the next step towards a working reactor, with the goal of producing a fusion/heating power ratio of $Q \sim 10$.

JET has a large available external heating power provided by Joule heating of the main solenoid current drive (*Ohmic*), *Neutral Beam Injection (NBI)*, *Ion Cyclotron Resonance Heating (ICRH)* (also known as Radio Frequency (RF) heating) and by Lower Hybrid Current Driving (LHCD). Early in its operational life, it formed a simple limited plasma and was able to produce plasmas with volumes of $\sim 200m^3$. Later, a pumped divertor was built to reduce impurities and aid establishment of the ETB and H-Mode operation but reduced the plasma volume to $\sim 90m^3$. Figure 1.5 shows the poloidal cross-section of JET during these two phases along with typical flux-surfaces. Table 1.2 gives the main operational ranges and the order of magnitude of some relevant and important plasma parameters.

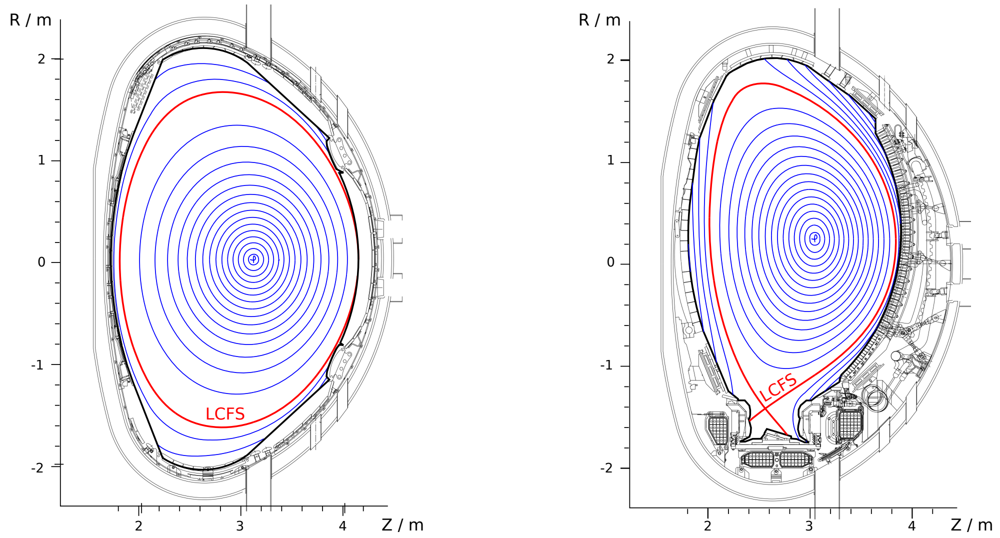

Figure 1.5.: left) Early and right) more recent view of poloidal cross-section of JET, showing vessel structure and magnetic flux surfaces.

B_ϕ	Vacuum Toroidal Magnetic Field (max at centre)	$3.5\,T$
I_p	Plasma Current (max)	$5\,MA$
P_{tot}	Total Aux. Heating Power (max)	$30\,MW$
β	Plasma Pressure / Magnetic Field (max)	$\sim 4\%$
n_e	Electron Density	$\sim 10^{20}\ m^{-3}$
T_e	Electron Temperature	$\sim 1-10\,keV$
T_i	Ion Temperature	$\sim 1-50\,keV$
c_a	Alfen speed	$\sim 10^7\ ms^{-1}$
c_s	Sound speed	$\sim 10^6\ ms^{-1}$
v_ϕ	Toroidal Velocity	$\sim 10^5\ ms^{-1}$
v_θ	Poloidal Velocity	$\sim 10^3\ ms^{-1}$
ρ_i	Ion Lamour Radius	$\sim 10^{-3}\ m$
ρ_e	Electron Lamour Radius	$\sim 10^{-4}\ m$
λ_d	Debye Length	$\sim 10^{-5}\ m$
ω_{pe}	Electron Plasma Frequency	$\sim 10^{11}\ Hz$
ω_{ce}	Electron Cyclotron Frequency	$\sim 10^{11}\ Hz$
ν_e	Electron-Electron Collision Frequency	$\sim 10^5\ Hz$
ν_i	Ion-Ion Collision Frequency	$\sim 10^4\ Hz$
ν_i	Electron-Ion Collision Frequency	$\sim 10^4\ Hz$

Table 1.2.: Capabilities and typical conditions for JET plasmas

1.6. JET Diagnostics

This sections gives an introduction to several of the plasma diagnostics used at JET. The modelling of three of these is the principal concern of this work and they are covered in much greater detail in later sections. The others are used for comparison, isolation of interesting pulses and time periods or to highlight plasma physics effects. While many of the details here are specific to the apparatus set up on JET, variations of all these systems can be found on many other Tokamaks where the physical principles of the measurements remain the same.

Figure 1.6 shows the lines of sight along which, or positions at which each diagnostic makes its measurement in the poloidal plane of JET.

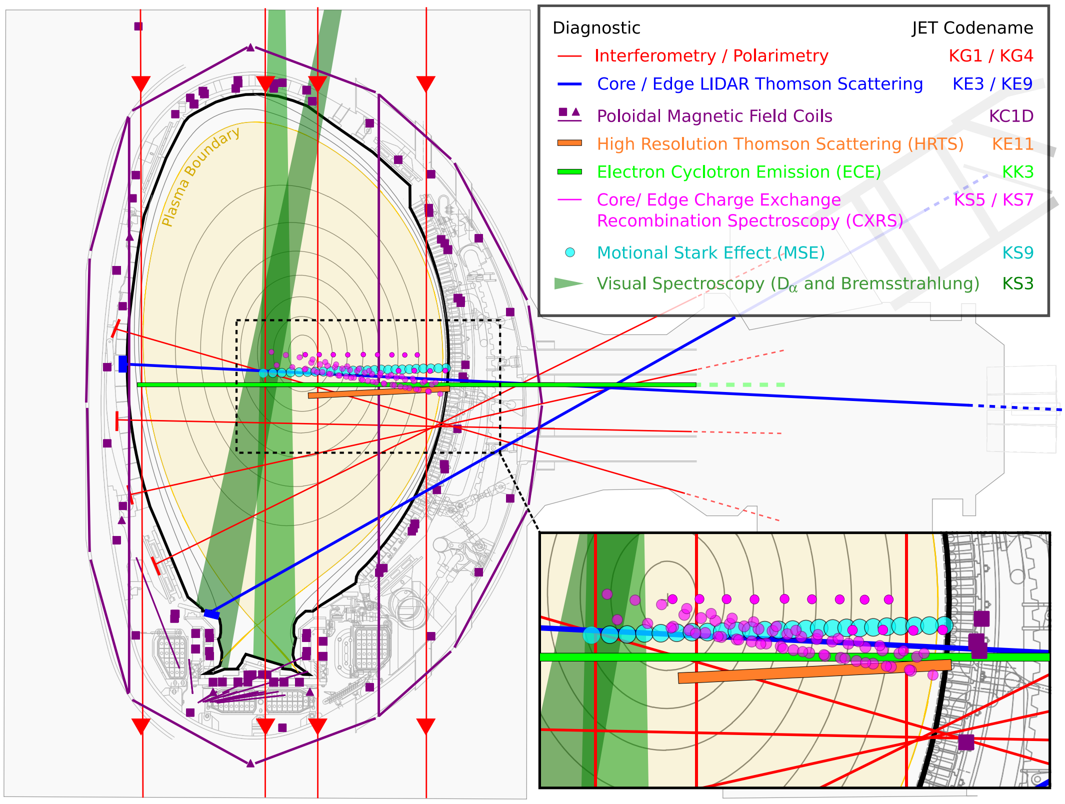

Figure 1.6.: The lines of sight or observation position(s) for all diagnostics on JET which are referenced in this work. These are shown projected into the poloidal plane along with a set of flux surfaces and boundary typical of a JET diverted H-Mode plasma. The first wall is shown in black.

Each diagnostic has its own independent *standard analysis* procedure and code that derives estimates of quantities of interest from the observed data. These are sometimes simple and general but in many cases, in order to be robust, make strong and sometimes inaccurate assumptions about the plasma and/or diagnostic behaviour. The standard analysis results are used for comparison and discussion throughout this document but the models, analysis and hence the principal results of this work are based on the raw observed data and so are entirely independent of these.

1.6.1. Modelled diagnostics

Interferometry

The interferometry system determines the line-integrated electron density n_e along 4 vertical and 4 lateral lines of sight. A far-infrared laser beam is sent along each path and interfered with a reference beam to measure the optical path difference from which the line integrated density is inferred. The noise level is very low and the standard analysis results are quoted as accurate to $10^{17}m^{-2}$, typically $< 1\%$. However, at any time, the system gives the offset from an unknown integer multiple of $1.143 \times 10^{19}m^{-2}$, called *fringes*. The current fringe is known only by following the evolution in time and rapid changes can result in 'fringe jumps' where the absolute value is lost and only the offset known. Such cases must be manually corrected, often by referring to other diagnostics.

Polarimetry

The polarimetry system measures the polarisation change of the interferometry laser beams due to the plasma and hence much of the hardware is shared by the two systems. Using a simple approximation, the standard analysis code gives the Faraday rotation $\Delta\psi \sim \int n_e(l)B_{\|}(l)\, dl$ and the induced ellipticity due to the Cotton-Mouton effect which follows $\tan\chi \sim \int n_e(l)B_{\phi}(l)\, dl$ if appropriately set-up. The Faraday rotation can be of great use in the inference of the magnetic topology since it gives measurements of the field induced by the plasma current, inside the plasma.

LIDAR/Thomson Scattering

Thomson scattering diagnostics infer the electron temperature T_e and density n_e from the intensity and Doppler broadening of Thomson scattered light, originating from a laser pulse sent into the plasma. For LIDAR systems, 180^o back-scattering is collected by a single spectrometer and the time of flight used to give complete profiles along the laser path. JET has two systems which both use a $\sim 300ps$ ruby laser pulse, one passing through the plasma core and one skimming the plasma edge. The core system fires once every $250ms$ throughout each shot and the standard analysis gives a spatial resolution of $\sim 12cm$. The edge system fires 6 times per pulse at $1s$ intervals and has a spatial resolution only slightly higher along its line of sight but when mapped on flux-surfaces to the mid-plane, this is $\sim 2cm$ using the standard analysis code. Both systems are quite complex and involve a large number of calibration parameters. Uncertainty in these and in the mapping causes significant problems in the use of profiles inferred using the standard analysis codes.

1.6.2. Other Diagnostics

Poloidal Magnetic Field Coils

The JET magnetic diagnostics set measures the poloidal magnetic field and flux outside the plasma and consists of 230 pickup coils, 88 saddle coils and 6 full toroidal flux loops. They are used for MHD mode analysis but primarily act as constraints to solve the Grad-Shafranov equation in order to infer the magnetic topology and plasma current profile. The JET equilibrium fitting code (EFITJ) runs as a standard analysis code using only the magnetics coils as constraints and using a heavily constrained current profile parameterisation to provide an approximate but robust result. More accurate results can be obtained with manual EFITJ runs, using internal measurements from other diagnostics, and weaker profile constraints[25].

Motional Stark Effect (MSE)

The JET motional stark effect diagnostic measures the Stark splitting of the D_α line emission from neutral beam particles. The emission is Doppler shifted due to the high beam particle velocity so can be separated from the background D_α emission. The line is split by the electric field in the particle's rest frame, which gives some information on the magnetic field in the lab frame. This can be used as one of the extra internal constraints for manual EFITJ runs[25]. The diagnostic data is not used directly and the system is only defined here as it is discussed on a few occasions in this work.

High Resolution Thomson Scattering (HRTS)

The high resolution Thomson scattering [26] is a conventional TS system with a series of separate spectrometers each viewing the laser propagation path at $\approx 90^o$. Installation of the system was completed about half-way through this project so the diagnostic is used only for comparison with the main results. The resolution (both integration length and channel spacing) of the system is $\approx 3cm$ for most of the JET pulses which were available and $\approx 1.5cm$ for a small amount of very recent data and the system had a $50Hz$ temporal sampling rate throughout.

Unlike the LIDAR systems, each channel must be calibrated separately and the complexity was such that all of the HRTS data available during this project was inferred using calibration coefficients determined by statistical comparison with other diagnostics. Both the absolute magnitude and profile shape are effectively determined by this.

Electron Cyclotron Emission (ECE)

In conditions usually true in JET plasmas, the intensity of electron-cyclotron emission is dependent on the local electron temperature and the frequency is a harmonic of the local electron cyclotron frequency, which depends only on the magnetic field. The field is dominated by the vacuum toroidal field which simply falls as $1/R$, giving a simple relationship between position and frequency.

Two systems on JET measure ECE radiation leaving the plasma [27, 28]. The first is a Michelson interferometer which has a low temporal resolution (15ms) but is absolutely calibrated. The second is a heterodyne radiometer [29] which gives high temporal resolution ($\sim \mu s$) and a good spatial resolution ($\sim 3cm$) but is cross-calibrated against the Michelson system. The heterodyne radiometer is also only able to infer T_e out to part way through the ETB, because the plasma becomes optically thin and the assumptions used by the standard analysis code break down. The standard analysis also produces T_e profiles which appear to be shifted by $\sim 3cm$ relative to the HRTS and LIDAR standard analysis results. Forward modelling of the ECE system is being carried out elsewhere at the time of writing and once complete, combination of this with the models and techniques developed here will be simple. Here, the standard analysis T_e results are used only for comparison.

Charge Exchange Recombination Spectroscopy (CXRS)

The Doppler shift and broadening of the emission created during charge exchange between charged impurity ions and injected neutral beam ions give information about the temperature and velocity of the impurity ions. Assuming that these are in thermal equilibrium with the bulk plasma ions, these allow the inference of the plasma ion temperature T_i and rotational velocity ω_i. It is also possible to infer the density of the impurity ions from their spectral lines and make an estimate of Z_{eff}, the effective charge of the plasma.

JET has two main charge exchange recombination spectroscopy (CXRS) diagnostics, one covering the plasma core[30] and other concentrating on the plasma edge[31].

Visible Spectroscopy (D-Alpha)

The JET visible spectroscopy diagnostic is used to measure the D_α emission from deuterium atoms and to estimate Z_{eff} from Bremsstrahlung emission. The D_α signal relates to the *recycling* - the flux of deuterium ions leaving the plasma, neutralising on or near the walls and re-entering the plasma. This allows clear identification of the start of H-Mode as the formation of the ETB reduces the recycling and the D_α falls below the L-Mode level. ELMs can also be seen clearly as a very short peak in the D_α signal, as was shown in section 1.4.4.

1.7. Motivation - Tokamak Data Analysis

To achieve the conditions for fusion, the plasma must be carefully controlled. This requires accurate knowledge of many aspects of the plasma state and the ability to predict its behaviour. Such understanding comes either empirically from experimental observations or from theoretical models, which must be checked against observations. The accuracy of the observations required varies but in every case, the accuracy achieved - the uncertainty - must be known.

1.7.1. Complementary information

The wide range of diagnostics on modern Tokamaks means that a large amount of information is available but also highlights problems. For example, where multiple diagnostics are used to infer the same physical quantity, they often conflict or disagree. This might be due to inaccuracies in calibrations or inconsistencies in the assumptions made about the diagnostic and/or plasma physics during the analysis of each data set. Handling the issues and complexity involved is a problem common to many experimental sciences, and is a field of study in its own right.

Under standard approaches, the primary purpose of each diagnostic is seen as the measurement of a single or a small number of physical parameters but calculating them often requires other parameters. Physical assumptions can also require other information, such as assuming quantities are constant on flux surfaces requires the magnetic topology to be known. Seen as inconvenient dependencies, they are taken from the analysis of another diagnostic which may have involved many assumptions itself. This leads to a chain of collected uncertainties which in many cases are too complex to calculate and propagate, so are omitted or crudely estimated. The lack of calculated uncertainties in flux surfaces from equilibrium fitting codes is a specific problem which has been particularly common in almost all Tokamak analysis for a long time, and one which this project attempts to address.

These dependencies can be seen as a benefit instead of an unavoidable complication, as they imply that the diagnostic holds some information about that parameter which can supplement that from the other source. However, to consistently combine multiple systems requires a way to treat the diagnostics simultaneously, with a single set of physical assumptions.

1.7.2. Low information content

In other cases, the information available can be extremely limited. Low signal to noise or large calibration uncertainty often gives such poor results through the standard analysis that data is entirely rejected. With careful handling of the uncertainties, information could be extracted from such data, especially when large quantities of it are available. JET has been in operation for 27 years and many of the diagnostics (especially those modelled in this work) have been recording data regularly for much of this time. In many fields of scientific research, physics effects are extracted from signals well beneath the noise level, by collecting large quantities of data but this is rarely done in Tokamak plasma analysis.

To really achieve the best possible use of the vast quantity of data collected by Tokamaks like JET, requires methods that can handle the high level of complexity that comes from combining a large number of very different measurement systems. These methods must rigorously keep track of all types of uncertainty in all measurements, without loss of any useful information that they may contain. The following chapter introduces both the mathematical concepts and computational methods that allow this to be achieved.

2. Analysis Techniques

2.1. Forward Modelling and Joint Analysis of Diagnostics

Even for individual diagnostics, calculating physical parameters from the observed data - the 'inverse' problem, is often difficult or impossible because many physical states might create similar or identical data. However, since the physics and operation of a diagnostic system are usually well understood, it is normally possible (though not necessarily trivial) to calculate a prediction of the data that would be produced, given a specified physical state. This is known as the *forward model* and representing the diagnostic systems in this way leads to a conceptually clean and simple way of dealing with most of the problems discussed in section 1.7.

The forward model clearly defines the dependencies of the diagnostic as seen in figure 2.1 which shows a (very simplified) forward model for a 3-channel single-point Thomson Scattering (TS) diagnostic, requiring the electron density n_e and temperature T_e as a function of space.

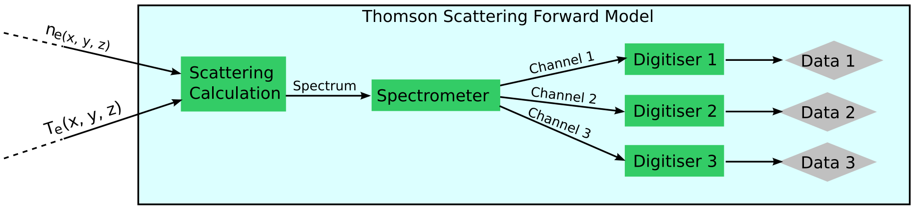

Figure 2.1.: Schematic of a simplified 3-channel single point Thomson Scattering diagnostic forward model, depending on electron density n_e and temperature T_e.

The dependencies can be tied to any plasma physics model/parameterisation, which is separate from the diagnostic model. Adding further diagnostics is simply a case of tying them to the same physics model. Figure 2.2 shows a simple combination of the TS system, an interferometer and some magnetic coils, tied to an axisymmetric model of the plasma current and electron density and temperature, which are assumed constant on flux surfaces.

The natural modularity lends itself to the implementation and this concept underpins the software framework used for this work - Minerva[32]. Each diagnostic model can be an independent computer code module and complete models built by linking together any combination of these with the desired physics model. The parameters can then be set and each module produces the predicted data for its diagnostic. Large and very complex systems can be assembled easily because, other than its dependencies, the details of the diagnostic model need not be known by the modeller. There are in principle no consistency issues with assumption since the forward models make no assumptions about the plasma physics (besides what is obvious from their dependencies, e.g. the dependency on T_e inherently assumes a Maxwellian velocity distribution).

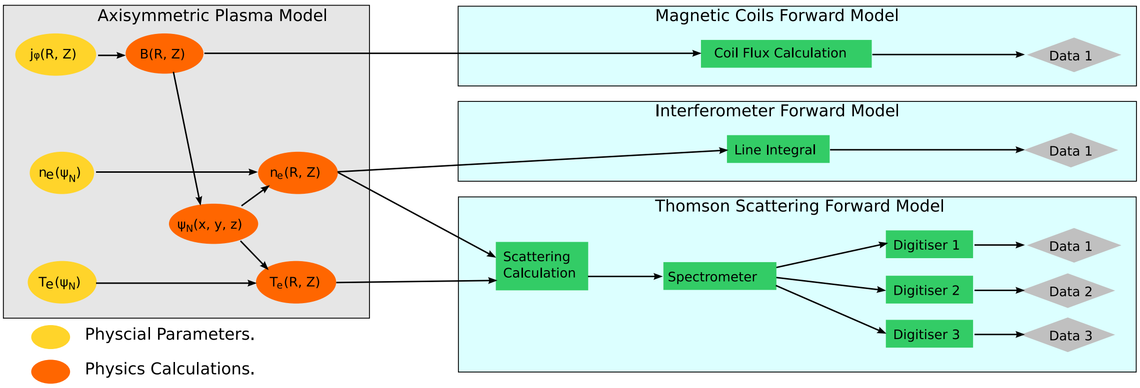

Figure 2.2.: Schematic of a simplified 3-diagnostic system, connected to an axisymmetric plasma model with parameterised $j_\phi(R, Z)$ and electron density and temperature modelled as 1D functions of normalised flux: $n_e(\psi_N), T_e(\psi_N)$.

An obvious side-effect is that the forward models can be used for diagnostic simulation but the real desire is to determine the parameters, given the measured data. In a simplified sense, this is just a case of searching for every set of the parameters for which the modules predict data very close to what was recorded. Doing this rigorously and handling the uncertainties, requires a framework for dealing with problem of inference itself.

2.2. Bayesian Analysis

2.2.1. The Likelihood Distribution

Ideally, the diagnostic models would provide the exact *forward function* $\underline{\mathrm{D}} = f_f\left(\underline{\mu}\right)$, which gives the vector of precise data that would be observed $\underline{\mathrm{D}}$ if the state of the entire system were exactly that described by the vector of parameters $\underline{\mu}$. This is never practically possible since there are always some details of the physical state too subtle or impossible to assign parameters to, such as electronic noise created from thermal fluctuations or noise arising from quantum effects which are inherently random. These contributions can instead be assigned a *Probability Density Function (PDF).* Instead of $f_f\left(\underline{\mu}\right)$, the diagnostic model produces a PDF of the data that would be observed $P\left(\underline{\mathrm{D}} \mid \underline{\mu}\right)$, known as the *likelihood distribution*, in repeated experiments under exactly the same conditions for the parameterised part of the physical state $\underline{\mu}$.

2.2.2. The Posterior Distribution and Bayes Theorem

The likelihood distribution still describes the forward situation, as it gives the uncertainty in $\underline{\mathrm{D}}$, given an exact $\underline{\mu}$. For a real experiment (the inverse/inference problem), the data $\underline{\mathrm{D}}$ observed is known exactly and the true physical state $\underline{\mu}$ is uncertain or not known. The PDF desired is $P(\underline{\mu} \mid \underline{\mathrm{D}})$ and is known as the *posterior* PDF. The posterior describes the probability that the system was in any given state $\underline{\mu}$ given that the data $\underline{\mathrm{D}}$ was observed and it can be obtained from the likelihood through the product rule of probability theory:

$$P\left(A,B\right) = P\left(A|B\right)\ P\left(B\right) \tag{2.1}$$

$$= P\left(B|A\right)\ P\left(A\right) \tag{2.2}$$

The two right hand expressions can be easily rearranged to the required one, known as Bayes Theorem:

$$P\left(A|B\right) = \frac{P\left(B|A\right)\ P\left(A\right)}{P\left(B\right)} \tag{2.3}$$

In terms of the plasma state and measured data, this is:

$$P\left(\underline{\mu} \mid \underline{\mathrm{D}}\right) = \frac{P\left(\underline{\mathrm{D}} \mid \underline{\mu}\right)\ P\left(\underline{\mu}\right)}{P\left(\underline{\mathrm{D}}\right)} \tag{2.4}$$

Since $\underline{\mathrm{D}}$ is known exactly, $P\left(\underline{\mathrm{D}}\right)$ which is known as the evidence and is a function of D alone, is constant. After dropping this by rewriting the equation as a proportionality, it contains only the posterior, likelihood and $P\left(\underline{\mu}\right)$ which is known as the *prior* probability. The prior encodes any knowledge, or lack of knowledge, of the plasma state before the measurement took place.

The posterior represents everything that is known about the plasma state from the diagnostic's observations and the prior assumptions, including all uncertainty. It always exists, and because it can represent inferred relationships between parameters, it can be calculated even for diagnostics that would normally be considered entirely dependent on another. Take for example, a diagnostic that really measures the data $D = \mu_1 + \mu_2 + \delta$, where δ is some small noise source. This might normally be considered a device for measuring μ_1 which is dependant on a separate measurement of μ_2. The posterior for the first diagnostic alone, would have a narrow ridge of high probability in the 2D plane of $P(\mu_1, \mu_2 | D)$, consistent with the measured sum.

Bayes Theorem and the techniques of Bayesian analysis have been used previously in nuclear fusion research [32, 33, 34, 35, 36, 37, 38].

2.2.3. The Joint Posterior Distribution

For models with multiple diagnostics, each forward model gives a separate likelihood function, based on the same plasma $P\left(\underline{\mathrm{D_1}} \mid \underline{\mu}\right)$, $P\left(\underline{\mathrm{D_2}} \mid \underline{\mu}\right)$ etc. Probability theory provides a way to handle models with multiple diagnostics and determine what can be inferred about $\underline{\mu}$ given all the data - the joint posterior $P\left(\mu \mid D_1, D_2\right)$. Starting with Bayes Theorem:

$$P\left(\underline{\mu} \mid \underline{\mathrm{D_1}}, \underline{\mathrm{D_2}}\right) \propto P\left(\underline{\mathrm{D_1}}, \underline{\mathrm{D_2}} \mid \underline{\mu}\right) \;\; P\left(\underline{\mu}\right) \tag{2.5}$$

The physical state common to more than one diagnostic should be held in $\underline{\mu}$, so that for fixed $\underline{\mu}$, variation in the data of each diagnostic - the noise - is independent of all other diagnostics. This makes the likelihood distributions entirely independent, and the joint posterior is simply:

$$P\left(\underline{\mu} \mid \underline{\mathrm{D_1}}, \underline{\mathrm{D_2}}\right) \propto P\left(\underline{\mathrm{D_1}} \mid \underline{\mu}\right) \;\; P\left(\underline{\mathrm{D_2}} \mid \underline{\mu}\right) \;\; P\left(\underline{\mu}\right) \tag{2.6}$$

2.2.4. Marginals and Conditionals

If a single parameter, or a subset of the parameters in $\underline{\mu}$ are fixed at some specified value, the PDF over the remaining parameter(s) is known as the *conditional* distribution, e.g. $P\left(\mu_1 \mid \underline{\mathrm{D}}, \mu_2, \mu_3, \ldots \mu_N\right)$. It is usually trivial to calculate and describes what would be known about μ_1, if the others are exactly that specified. In most circumstances, what is desired is what can be inferred about the subset, independent of all the others e.g. $P\left(\mu_1 \mid \underline{\mathrm{D}}\right)$. This is known as the *marginal* distribution and is significantly more complex to calculate, as it requires integrating over the other parameters:

$$P\left(\mu_1 \,|\, \underline{\mathrm{D}}\right) = \int_{-\infty}^{\infty} P\left(\underline{\mu} \,|\, \underline{\mathrm{D}}\right)\, d\mu_2\, d\mu_3\, \ldots\, d\mu_N \tag{2.7}$$

2.2.5. Bayesian Analysis In Practice

For most problems $\underline{\mu}$ will include a very large number of parameters. For instance both $\underline{\mathrm{n_e}}$ and $\underline{\mathrm{T_e}}$ might be parameterised as 1D profiles through an interpolation of 100 nodes each, making the posterior $P(\underline{\mu} \,|\, \underline{\mathrm{D}})$ a 200-dimensional object and impossible to calculate, store or interpret. In some special circumstances, it can be entirely described by low order moments, for example when the posterior is a multivariate Gaussian, the 200 element mean and 200 × 200 covariance matrix is sufficient. However, the posterior is usually represented by a series of *samples* of $\underline{\mu}$, calculated to be as if they were representative samples drawn from the posterior. Also, the $\underline{\mu}$ with the highest posterior probability density - the *Maximum posterior* (MAP) can be found which is effectively the 'best fit' of traditional methods. It should be noted however, that the MAP is not necessarily close to the mean posterior and that the value for a single parameter at the MAP, might not be close to the value of that parameter at the maximum of its marginal distribution, which is the best estimate for that parameter independently of the others. A benefit of the storage of the posterior as samples, is that the distribution of a set of parameters within them, is that parameter's marginal distribution. There are no further steps required to calculate the high-dimensionality integral of equation 2.7.

2.2.6. Inversion Algorithms and Minerva

The methods used to find the MAP and generate samples are covered in 2.3 and 2.4, but it is a particular feature of the analysis that the solution methods are entirely external to the model, both conceptually and in the implementation. This is unlike many of the standard analysis codes which perform fits, estimations and other inversion methods internally at multiple stages in the analysis chain.

All of the work described in this thesis was carried out within *Minerva*[32], a software framework for handling high complexity forward modelled Bayesian inference problems, described in terms of a Bayesian Graphical Models[4]. As well as the diagnostic models, work by the author for this thesis included development of common physics modules, inversion algorithms and a general parallelisation suite, as well as some minor work on the framework itself. All of these are completely reusable modules, available to anyone working with Minerva and some of this is already in use by other parties both at JET and on other Tokamaks such as the *Mega-Amp Spherical Tokamak (MAST)* [39, 40].

2.2.7. Bayesian Theory and Terminology in this thesis

The Bayesian theory set out above provides a rigorous basis for the analysis carried out in this work. However, since the primary objective of the thesis is the analysis of Tokamak plasmas, it is used as tool. In some cases, more traditional methods (like least-squares fits) are simpler to perform and describe and can be proven to follow from Bayesian theory under certain assumptions. In other cases, traditional language is used simply as it is shorter and will be more familiar to most readers. As such, the principles, precise mathematical formalism and especially the terminology of Bayesian theory are not followed exactly in every part. More detailed and careful introductions to Bayesian analysis methods are available elsewhere[41].

2.3. The Linear Gaussian Inversion

The *linear Gaussian inversion (LGI)* is an extremely useful numerical tool used in a specific situation to calculate the posterior distribution in a single relatively simple operation (from [38] and [42] before that). Although occurrences of the exact conditions are rare, many problems can be well approximated in a compatible way.

The method is applicable when the likelihood of the data vector $\underline{\mathrm{D}}$ can be expressed as a multivariate Gaussian with some known covariance $\underline{\underline{\sigma_{\mathrm{D}}}}$, and that the mean $\underline{\mathrm{D}_0}$ is a known linear combination of the unknown parameters $\underline{\mathrm{D}_0} = \underline{\underline{\mathrm{M}}}\,\underline{\mu} + \underline{\mathrm{C}}$. The *response matrix* $\underline{\underline{\mathrm{M}}}$ and the constants $\underline{\mathrm{C}}$ make up the linear forward function in this case. The final requirement is that the prior must also be expressed as multivariate Gaussian with mean $\underline{\mu_{\mathrm{P}}}$ and covariance $\underline{\underline{\sigma_{\mathrm{P}}}}$:

$$\begin{aligned} P\left(\underline{\mathrm{D}} \,|\, \underline{\mu}\right) &= \mathcal{G}\left(\underline{\mathrm{D}};\ \underline{\underline{\mathrm{M}}}\,\underline{\mu} + \underline{\mathrm{C}},\ \underline{\underline{\sigma_{\mathrm{D}}}}\right) && (2.8) \\ P\left(\underline{\mu}\right) &= \mathcal{G}\left(\underline{\mu};\ \underline{\mu_{\mathrm{P}}},\ \underline{\underline{\sigma_{\mathrm{P}}}}\right) && (2.9) \end{aligned}$$

Through Bayes theorem, the posterior, with an arbitrary normalisation constant A_1 is:

$$\begin{aligned} P\left(\underline{\mu} \,|\, \underline{\mathrm{D}}\right) &\propto P\left(\underline{\mathrm{D}} \,|\, \underline{\mu}\right) P\left(\underline{\mu}\right) \\ &\propto \mathcal{G}\left(\underline{\mathrm{D}};\ \underline{\underline{\mathrm{M}}}\,\underline{\mu} + \underline{\mathrm{C}},\ \underline{\underline{\sigma_{\mathrm{D}}}}\right) \mathcal{G}\left(\underline{\mu};\ \underline{\mu_{\mathrm{P}}},\ \underline{\underline{\sigma_{\mathrm{P}}}}\right) \\ \log P\left(\underline{\mu} \,|\, \underline{\mathrm{D}}\right) &= -\tfrac{1}{2}\left(\underline{\mathrm{D}} - \underline{\underline{\mathrm{M}}}\,\underline{\mu} - \underline{\mathrm{C}}\right)^T \underline{\underline{\sigma_{\mathrm{D}}}}^{-1} \left(\underline{\mathrm{D}} - \underline{\underline{\mathrm{M}}}\,\underline{\mu} - \underline{\mathrm{C}}\right) && (2.10) \\ &\quad -\tfrac{1}{2}\left(\underline{\mu} - \underline{\mu_{\mathrm{P}}}\right)^T \underline{\underline{\sigma_{\mathrm{P}}}}^{-1} \left(\underline{\mu} - \underline{\mu_{\mathrm{P}}}\right) + A_1 \end{aligned}$$

Since the posterior is a multiplication of the two Gaussians, it can be written as a single Gaussian with mean $\underline{\mu_0}$ and covariance $\underline{\underline{\sigma}}$:

$$\begin{aligned} P\left(\underline{\mu} \,|\, \underline{\mathrm{D}}\right) &\propto \mathcal{G}\left(\underline{\mu};\ \underline{\mu_0},\ \underline{\underline{\sigma}}\right) \\ \log P\left(\underline{\mu} \,|\, \underline{\mathrm{D}}\right) &= -\tfrac{1}{2}\left(\underline{\mu} - \underline{\mu_0}\right)^T \underline{\underline{\sigma}}^{-1} \left(\underline{\mu} - \underline{\mu_0}\right) + A_2 && (2.11) \end{aligned}$$

Expanding equations 2.10 and 2.11 and equating like terms in $\underline{\mu}$:

$$\begin{aligned}
\underline{\mu}^T \underline{\underline{\sigma}}^{-1} \underline{\mu} &= \underline{\mu}^T \left[\underline{\underline{\mathrm{M}}}^T \underline{\underline{\sigma_{\mathrm{D}}}}^{-1} \underline{\underline{\mathrm{M}}} + \underline{\underline{\sigma_{\mathrm{P}}}}^{-1} \right] \underline{\mu} && (2.12) \\
\underline{\mu_0}^T \underline{\underline{\sigma}}^{-1} \underline{\mu} &= \left[\left(\underline{\mathrm{D}} - \underline{\mathrm{C}} \right)^T \underline{\underline{\sigma_{\mathrm{D}}}}^{-1} \underline{\underline{\mathrm{M}}} + \underline{\mu_{\mathrm{P}}}^T \underline{\underline{\sigma_{\mathrm{P}}}}^{-1} \right] \underline{\mu} && (2.13) \\
\underline{\mu}^T \underline{\underline{\sigma}}^{-1} \underline{\mu_0} &= \underline{\mu}^T \left[\underline{\underline{\mathrm{M}}}^T \underline{\underline{\sigma_{\mathrm{D}}}}^{-1} \left(\underline{\mathrm{D}} - \underline{\mathrm{C}} \right) + \underline{\underline{\sigma_{\mathrm{P}}}}^{-1} \underline{\mu_{\mathrm{P}}} \right] && (2.14) \\
\underline{\mu_0}^T \underline{\underline{\sigma}}^{-1} \underline{\mu_0} + A_2 &= \left(\underline{\mathrm{D}} - \underline{\mathrm{C}} \right)^T \underline{\underline{\sigma_{\mathrm{D}}}}^{-1} \left(\underline{\mathrm{D}} - \underline{\mathrm{C}} \right) - \tfrac{1}{2} \underline{\mu_{\mathrm{P}}}^T \underline{\underline{\sigma_{\mathrm{P}}}}^{-1} \underline{\mu_{\mathrm{P}}} + A_1 && (2.15)
\end{aligned}$$

From equations 2.13 and 2.14, the posterior covariance and mean are:

$$\begin{aligned}
\underline{\underline{\sigma}} &= \left[\underline{\underline{\mathrm{M}}}^T \underline{\underline{\sigma_{\mathrm{D}}}}^{-1} \underline{\underline{\mathrm{M}}} + \underline{\underline{\sigma_{\mathrm{P}}}}^{-1} \right]^{-1} && (2.16) \\
\underline{\mu_0} &= \underline{\underline{\sigma}} \left[\underline{\underline{\mathrm{M}}}^T \underline{\underline{\sigma_{\mathrm{D}}}}^{-1} \left(\underline{\mathrm{D}} - \underline{\mathrm{C}} \right) + \underline{\underline{\sigma_{\mathrm{P}}}}^{-1} \underline{\mu_{\mathrm{P}}} \right] && (2.17)
\end{aligned}$$

The power of the method can be seen in these last two equations - that the full mean and covariance of the posterior distribution can be calculated in a single matrix inversion given the inverse covariances of the prior and likelihood distributions and the response matrix.

In general (where the LGI conditions are not met), the posterior maximum is found by iterative numerical algorithms and the shape and extent investigated by the drawing random samples, a procedure which usually involves a gradual random walk. Both of these processes can take many thousands of times the parameter dimensionality of forward function evaluations to complete and for very high dimensionality ($N(\underline{\mu}) > \sim 200$) becomes prohibitively expensive.

For the LGI, the maximum posterior $\underline{\mu_0}$ is immediately available and the drawing of random samples from a multivariate Gaussian is a well known, relatively trivial procedure. Every marginal distribution is also Gaussian, with mean and covariance found simply by dropping the rows and columns of $\underline{\mu_0}$ and $\underline{\underline{\sigma}}$ that relate to the parameters to be marginalised out.

2.3.1. Practical Application and Parallel Implementation

Often, the forward function is easily written as a linear combination of the parameters and the coefficients are used to directly construct $\underline{\underline{\mathrm{M}}}$ and $\underline{\mathrm{C}}$. For more complex cases, they are found by first calculating the data $\underline{\mathrm{D}_{\mathrm{i}}}$ for some initial set of parameters $\underline{\mu_{\mathrm{i}}}$ about which the forward function is assumed linear. Each parameter in turn is then modified and the relevant parts of $\underline{\underline{\mathrm{M}}}$ and $\underline{\underline{\mathrm{C}}}$ filled using the difference between the new predicted data and $\underline{\mathrm{D}_{\mathrm{i}}}$. The calculation requires only the one evaluation of the forward function for each parameter and the initial one, so is relatively low cost. This proceedure can be performed automatically on any model within Minerva.

However, in some cases this work requires repeated application of the LGI to very large problems ($N(\underline{\mu}) \sim N(\underline{\mathrm{D}}) > 4000$) with relatively slow forward functions ($t \sim 100ms$). To reduce the inversion time, a general parallel implementation of the LGI was developed for Minerva. The determination of $\underline{\underline{\mathrm{M}}}$ is ideally suited to this since each machine can independently calculate a selection of the rows. Once complete, $\underline{\mathrm{M}}$ is distributed over the involved machines so that after distribution of $\underline{\underline{\sigma_{\mathrm{D}}}}^{-1}$, $\underline{\underline{\sigma_{\mathrm{P}}}}^{-1}$ and $\underline{\mu_{\mathrm{P}}}$, equations 2.16 and 2.17 are performed using freely available parallel matrix libraries (PBLAS, BLACS, scaLAPACK [43], [44]). The developed software allows the calculation of such large problems in a few minutes, where the serial implementation can take hours. It also allows the calculation of the posterior even when the model has too many parameters for the matricies to be held in the memory of a single computer, since the storage is distributed across the participating machines.

2.3.2. Truncated Gaussians

In many cases where the priors are not Gaussian, they can be easily represented as truncated Gaussians. Typical examples are densities and temperatures where parameters are restricted to being positive $P(n_e) = 0$ for $n_e < 0$. In such cases, the LGI procedure is applied as if the truncations were not present and the truncation simply applied directly to the posterior (There should also be a modification to the normalisation, but the correct normalisation of the posterior is rarely of any practical use).

If the Gaussian centre lies inside the truncation limits so $P(\underline{\mu_0}) \neq 0$ then it is the posterior maximum, otherwise the maximum will lie somewhere on one of the truncation hyper-planes and must be found by one of the general iterative algorithms.

The best known way of drawing random samples from the truncated posterior is a Monte-Carlo process, so is much slower than in the standard LGI, but it is much less costly than the general Monte-Carlo methods. The procedure is based on the Gibbs-sampler[45], where a random sample is drawn from the conditional distribution over one parameter given the current position in all others $P(\mu_i \mid \mu_{j\neq i})$. The sampler moves to the position in μ_i of the sample and the procedure is repeated for each parameter in turn. The whole process repeated many times so that position, moving in steps along each axis, gradually explores the entire joint distribution.

The procedure is favourable for the truncated multivariate Gaussian because the conditional distributions are always truncated univariate Gaussians whose mean and variance are easily calculated from $\underline{\mu_0}$ and $\underline{\underline{\sigma}}$ and methods exist to efficiently draw samples from these[46], so the overall procedure is fairly efficient. While this has been done previously[47], the posteriors in this work are often very highly correlated making the parameter-by-parameter Gibbs sampling relatively slow to move along the correlations. To mitigate this, a new Gibbs-sampler was developed where each step is taken along the next eigenvector of $\underline{\underline{\sigma}}$. The variance of the conditional along this line (still univariate Gaussian) is given by the associated eigenvalue. The distance along the line where it intersects the truncation hyper-plane on each real parameter is found and the most constrictive of these give the limits on that conditional. For the situations where it is used in this work, the procedure is significantly more efficient than that in [47].

The principle behind both the existing (parameter-space aligned) and the eigenvector aligned methods are shown in figure 2.3.

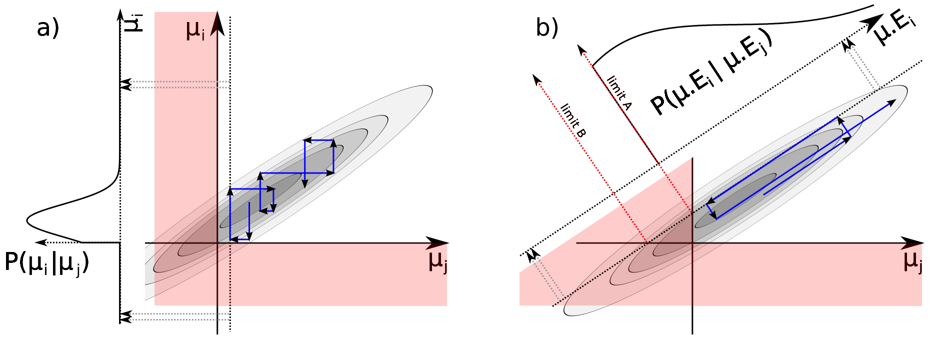

Figure 2.3.: The principle of the Gibbs-sampler applied to a 2D truncated Gaussian using a) parameter-space aligned conditionals/steps as in [47] and b) using eigenvector aligned conditionals as in this work. The green arrowed lines show some arbitrary example jumps and the 1D graphs show the conditional PDFs for one them in each case. $\underline{\mu}$ are the parameters and $\underline{\underline{E}}$ the eigenvectors of the Gaussian covariance.

2.4. Non-Linear Algorithms

Where the (T)LGI is not applicable, a number of different algorithms are available. These are split into those that find the position of the MAP (MAximum posterior) or a point close to it and those that explore the posterior distribution, producing a series of positions which are effectively samples of the posterior. The details are included here because they are an important part of the practical Bayesian methods, were used for most of results reported and because a large part of the project time was spent investigating, implementing, optimising and further developing them. However, the details should not effect the MAP or samples obtained in any way, so the casual reader may skip this section if willing to accept that the algorithms perform correctly.

2.4.1. Gradient and Line Search Algorithms

These are useful for cases where the posterior has a fairly low dimensionality, is smooth and has only a single maximum. They each involve first choosing a direction in parameter space and then finding the optimum position along that direction.

The line search can be done with robust methods like a bisection or preferably Golden Section search from the edges of parameter space or in some cases Newton Raphson iterations can be used. For posteriors with low correlation between parameters, taking the directions along each parameter in turn can be sufficient but the method more usually used is to choose the direction of steepest gradient at each stage. A more optimised version of this is the 'Conjugate Gradient' method [48].

Unfortunately, the best of these methods require the posterior gradients. If the posterior calculation is itself complex, as is usally the case, the gradients must be calculated numerically and correctly choosing the 'small' step size involved is, in general, difficult. Almost all interesting cases in this work were found to be far too complex, especially since the truncation of many parameters (e.g $P(n_e) = 0$ for $n_e < 0$) give infinite gradients which cannot be easily handled.

2.4.2. Pattern Search

Several algorithms, such as the 'Downhill simplex search' and 'Hooke and Jeeves' were tried, which perform pattern based searches through the whole parameter space, converging on the maximum. The advantage over the first class is that they do not require external gradient calculations so require less tuning. The Hook and Jeeves algorithm was found to be of use in many of the lower complexity cases in this work although the convergence time is usually long. Like the gradient algorithms, these require the space to be uni-modal or for the start position to be nearest to the global maxima.

2.4.3. The Genetic Algorithm

In most of the cases in this work, the MAP is found using a *Genetic Algorithm (GA)*. The GA can search extremely complex, very high dimensional posteriors with multiple maxima. It is a computational implementation of natural selection.

Basic Algorithm

The GA holds a collection of positions in parameter space, known as its *population* and generates new *child* solutions, either by *mutation* where an existing population member is taken and its parameters modified slightly or by *cross-over* where parameters are taken from two existing population members. The *fitness* ($\log P$ here) for the child is calculated and compared to an existing population member. The child can be either rejected or accepted (replacing that existing member) based on a random decision biased by the comparison. This process allows the GA to simultaneously search multiple 'good areas' of the posterior and slowly approach the optimum of each. When the process is deemed complete, the population member with the highest fitness is selected as the MAP solution.

The algorithm has a considerable degree of flexibility and hence can be heavily tuned to each problem by varying such things as the population size, the relative number of cross-overs and mutations, the exact details of the cross-overs and mutations or the behaviour and biasing involved in accepting or rejecting a new child.

Development and Parallelisation

Special thanks and recognition are extended to Dr. Alex Meakins who, with some assistance from the author, developed the GA for this project which includes special features such as an adaptive mutation system, which gradually adjusts the mutations attempted based on the acceptance rates of the previous mutations, greatly improving the convergence speed of the system and making many of the highest complexity problems later in this document possible within a reasonable time.

The GA was also parallelised by the author under a scheme where multiple GAs run on different computational units, each evaluating the posterior and progressing independently until a random selection of the population is exchanged with another GA in the cluster.

2.4.4. The Gibbs Sampler

The Gibbs sampler is used to explore the posterior (draw samples) in circumstances where samples can be easily drawn from the conditional distribution over a subset of parameters (possibly a single one) by another means. It is only used in this work as part of the TLGI procedure covered in the previous section (2.3.2).

2.4.5. The Metropolis Hastings Markov Chain Monte Carlo Sampler

Once the maximum is found, the *Markov Chain Monte Carlo (MCMC)* processes allow representative samples to be drawn the PDF. The algorithm used here is based on the Metropolis Hastings [49] algorithm.

The algorithm performs a random walk around the 'target' distribution (the posterior in this case). The individual *jumps* taken are sampled from from a *proposal* or *trial* distribution and then a decision of whether to proceed with the jump is made based on the ratio of the target probability at the existing and new positions. If the new position evaluates to a higher probability, the jump is made. If the new probability is lower, the decision to jump is taken randomly, biased by the probability ratio. If the jump is rejected, another sample is taken at the existing position before a new trial jump is drawn. The resulting series of positions is the Markov Chain and, given a long enough chain, can be shown to provide representative samples of the target distribution.

Burn-in, chain length and sampling efficiency

If the algorithm begins from a position too far from the MAP, the MCMC will eventually converge towards it (at least for a uni-modal distribution) but this may take a significant amount of time compared to the GA. If the MCMC begins exactly at the MAP, it will spend some time gradually moving away. This counter-intuitive effect is correct, since if true representative samples were drawn from the distribution, the total probability of a sample being drawn from the small region of high probability density around the MAP is extremely low, compared to the total probability of it being drawn from somewhere in the large (hyper)volume of low probability density further away. In either case, the chain must be allowed to complete this initial *burn-in* phase and reach a steady state before the samples can be used. Since the positions returned are separated only by small jumps, successive positions will be heavily correlated. To remove this artificial correlation, samples must be taken from a chain long enough to have explored the PDF sufficiently. If the proposal distribution is too small, this will take a very large number of jumps and hence a long time. If the proposal distribution is too large, too many attempted jumps will be rejected and the time before achieving a representative chain will also be too long. The optimum efficiency is obtained using a proposal distribution of the same shape as the target, scaled smaller by a fraction dictated by the dimensionality[50].

Adaptive Proposal

In low dimensionally or low correlation cases, using a small uncorrelated Gaussian for the proposal distribution is often sufficient. For the high correlation, low dimensionality cases many methods are available which gradually adapt the proposal distribution based on the samples obtained. A simple method is to use an uncorrelated Gaussian initially and after a short time after the burn-in has completed, calculating the sample covariance of the prematurely terminated chain. The chain is then resumed, using a Gaussian with the calculated covariance, rescaled according to the dimensionality. This was tested on a few known heavily correlated distributions and was able to successfully draw representative samples from them, up to a few hundred dimensions. It was used for many of the medium and high complexity cases in this work. For an in depth review and discussion of more rigorous adaptive MCMC algorithms, see [51].

LGI bootstrapped Metropolis Hastings

For very high dimensionality ($N > 500$), the number of samples required to obtain the covariance becomes prohibitively large and use of the MCMC becomes extremely difficult. However, in many of cases covered here, a large part of the parameters, data and model satisfy the conditions of the (T)LGI. While the approximation is not valid enough to use the LGI posterior itself, the covariance makes a sufficiently good proposal distribution, after the appropriate rescaling.

Parallelisation

If multiple independent MCMC chains are run in parallel, the burn-in time remains the same but once complete, each chain will begin exploring in a random direction. Samples taken from all chains will be representative long before those taken from an individual chain. Also, it is easier to determine if the chains have explored far enough, since the chains should cross each other in parameter space often.

2.5. Current Tomography

The first stage of the wider Bayesian analysis project at JET was developed prior to this work. It includes a full model of the JET poloidal magnetic diagnostics - flux loops, saddle and pick-up coils and their use to infer the PDF of possible toroidal plasma currents. The method, known as *Current Tomography*[38] (CT) is outlined here as it is the basis on which much of this work is built.

The toroidal current j_ϕ is modelled with uniform axisymmetric 'beams' of rectangular cross-section as shown in figure 2.4a. As well as the set inside the vessel used to parametrise the plasma current, a set is included to model the effect of the PF coils (for which the currents are known) and a further set to represent the unknown poloidal field contribution from the ferromagnetic iron core.

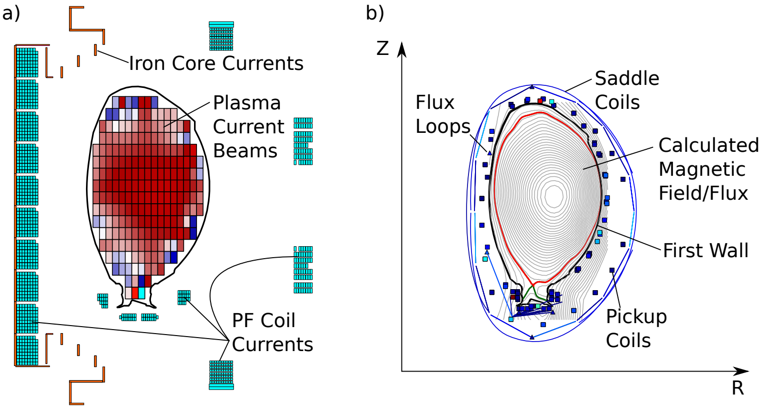

Figure 2.4.: JET Current Tomography: Poloidal cross-sections showing a) axisymmetric beams used to parametrise plasma, PF and iron core toroidal current and b) the poloidal magnetic flux and magnetic diagnostics in the forward model.

The physics forward model calculates B_θ and ψ and the diagnostic model calculates the predicted signal for the diagnostics coils shown in figure 2.4b. These are all linear functions of the current parameters so the LGI can be used to produce the full Gaussian posterior distribution of possible current configurations from the magnetics data D_m. Using only the magnetic diagnostics outside the plasma, the detail of the current distribution is very uncertain and the posterior is very wide (has large eigenvalues) in directions associated with short scale length variations. It is possible to add a 2D smoothing prior known as the *Conditional Auto-Regressive (CAR)* prior, representing a belief that neighbouring current beams should not vary significantly. With a weak CAR prior, the range of possible current distributions remains large but quantities calculated from them, such as the total toroidal current, magnetic axis position, strike points, flux surfaces etc. can be very well determined.

It should be noted that the method seeks to obtain the possible axisymmetric current distributions consistent with the measurements, with very weak assumptions about the form of the current distribution. It does not assume or involve equilibrium in any way and does not parametrise or involve the poloidal component of the plasma or coil currents.

2.5.1. Adding diagnostics and flux surface dependencies

In the original CT paper, the work was independent of an equilibrium assumption, and it was envisioned that the uncertainty should be reduced by adding diagnostic models and data, such as the MSE covered in the paper itself. In chapters 4 and 5 in this work, the JET Interferometry, Polarimetry and LIDAR Thomson Scattering diagnostic models are developmed, and added to the Current tomography model. These all either depend on j_ϕ directly, or do so via their dependence on n_e and T_e and the assumption that these are constant on each flux surface. When tied to the CT model, the CT flux surface uncertainty is effectively included in the analysis of those diagnostics and in some cases, the diagnostics provide some constraint on j_ϕ. Alternatively, the modelling and analysis of these diagnostics can be based on the fixed standard EFITJ solution or the fixed CT MAP result. In chapter 6, the flow-free equilibrium constraint is added directly to the CT system but without many of the other assumptions that are made in standard equilibrium solvers.

3. Interferometry And Polarimetry - Theory and Model

3.1. Electromagnetic Waves in Cold Plasma

This section gives a simplified derivation of the plasma wave propagation theory relevant to the diagnostic systems and models discussed in the remainder of the chapter. More rigorous derivations and more detailed discussions, applicable in much more general circumstances can be found elsewhere[52].

The diagnostic techniques modelled in this chapter involve the propagation through the plasma of an electromagnetic wave from a far-infrared laser. Such waves are relatively high frequency, low amplitude waves and for all JET plasmas, the laser frequency ω is far greater than both the electron plasma and electron cyclotron frequencies ω_p and ω_c and the wavelength far shorter than the scale of any large density variations. The low amplitude means that the wave can be modelled as a simple linear plane wave. The short wavelength allows the spatially varying plasma to be treated as independent small consecutive regions of homogeneous plasma, through which the evolution of the wave is modelled along a straight path, starting from the state it left the previous region (often referred to as the WKB approximation). The frequency is sufficiently high that the inertia of the plasma ions is too great for them to be effected significantly by the plasma wave and so their motion is neglected in this treatment. The currents induced by the wave's $\underline{\mathrm{E}}$ field is considered to be only movement in electrons.

A simple plane wave disturbance for this situation, travelling with wave vector $\underline{\mathrm{k}}$ and frequency ω can be expressed as a small oscillating electric field of magnitude $\underline{\mathrm{E}}_0$ and the accompanying movement of the plasma electrons, of magnitude $\underline{\mathrm{v}}_0$. These are given by the real part of:

$$\underline{\mathrm{E}} = \underline{\mathrm{E}}_0 e^{i(\underline{\mathrm{k}}\cdot\underline{\mathrm{r}}-\omega t)} \tag{3.1}$$

$$\underline{\mathrm{v}} = \underline{\mathrm{v}}_0 e^{i(\underline{\mathrm{k}}\cdot\underline{\mathrm{r}}-\omega t)} \tag{3.2}$$

More complex waves can be modelled simply as a linear sum of these.

The evolution of the electric and magnetic fields is described by the two Maxwell's equations:

$$\nabla\times\underline{\mathrm{E}} = -\frac{\partial\underline{\mathrm{B}}}{\partial t} \tag{3.3}$$

$$\nabla\times\underline{\mathrm{B}} = \mu_0\,\underline{\mathrm{j}} + \frac{1}{c^2}\frac{\partial\underline{\mathrm{E}}}{\partial t} \tag{3.4}$$

$$\tag{3.5}$$

To begin with, the wave is considered to be travelling in a vacuum and the plasma is included as the free current j, rather than considering it as a possibly dielectric/diamagnetic medium.

Taking the curl of the first and combining these gives the wave equation:

$$\nabla \times \nabla \times \underline{\mathrm{E}} = -\mu_0 \frac{\partial \underline{\mathrm{j}}}{\partial t} - \frac{1}{c^2}\frac{d^2 \underline{\mathrm{E}}}{dt^2} \tag{3.6}$$

The current is simply the movement of the electrons $\underline{\mathrm{j}} = -nq\underline{\mathrm{v}}$ where n and q are the electron density and charge. Substituting the plane wave solutions and taking derivatives:

$$\underline{\mathrm{k}} \times \underline{\mathrm{k}} \times \underline{\mathrm{E}} = i\omega nq\mu_0 \underline{\mathrm{v}} - \frac{\omega^2}{c^2}\underline{\mathrm{E}} \tag{3.7}$$

The next stage is to calculate the response of the plasma electrons to the wave by finding $\underline{\mathrm{v}}$ in terms of $\underline{\mathrm{E}}$. For a general case this is very complex but a simple derivation can be made using the electron fluid momentum equation, with the pressure and collisional terms ignored. This results in what is known as the *cold plasma approximation*:

$$mn\left[\frac{d\underline{\mathrm{v}}}{dt} + (\underline{\mathrm{v}} \cdot \nabla)\,\underline{\mathrm{v}}\right] = -nq\left(\underline{\mathrm{E}} + \underline{\mathrm{v}} \times \underline{\mathrm{B}}\right) \tag{3.8}$$

where m is the electron mass.

Substituting the plane waves solutions and dropping all terms above first order in disturbed quantities, as the disturbances are considered small, gives:

$$\underline{\mathrm{v}} = -\frac{iq}{m\omega}\left(\underline{\mathrm{E}} + \underline{\mathrm{v}} \times \underline{\mathrm{B}}\right) \tag{3.9}$$

$$\tag{3.10}$$

The dot and cross product of this with B are:

$$\underline{\mathrm{v}} \cdot \underline{\mathrm{B}} = -\frac{iq}{m\omega}\underline{\mathrm{E}} \cdot \underline{\mathrm{B}} \tag{3.11}$$

$$\underline{\mathrm{v}} \times \underline{\mathrm{B}} = -\frac{iq}{m\omega}\left(\underline{\mathrm{E}} \times \underline{\mathrm{B}} - B^2\underline{\mathrm{v}} + (\underline{\mathrm{B}} \cdot \underline{\mathrm{v}})\,\underline{\mathrm{B}}\right) \tag{3.12}$$

These can be substituted back into 3.9 and after splitting the magnetic field $\underline{\mathrm{B}} = B\underline{\hat{\mathrm{b}}}$ and defining the electron cyclotron frequency $\omega_c = qB/m$ gives the electron velocity response to the wave $\underline{\mathrm{E}}$ field:

$$\underline{\mathrm{v}} = \frac{-iq}{m\omega\left(1 - \frac{\omega_c^2}{\omega^2}\right)}\left[\underline{\mathrm{E}} + \frac{i\omega_c}{\omega}\underline{\mathrm{E}} \times \underline{\hat{\mathrm{b}}} - \frac{\omega_c^2}{\omega^2}\left(\underline{\mathrm{E}} \cdot \underline{\hat{\mathrm{b}}}\right)\underline{\hat{\mathrm{b}}}\right] \tag{3.13}$$

With this and the definition of electron plasma frequency $\omega_p^2 = nq^2/(\epsilon_0 m)$, the wave equation (3.7) becomes:

$$\frac{c^2}{\omega^2}\underline{\mathrm{k}} \times \underline{\mathrm{k}} \times \underline{\mathrm{E}} \quad = \quad \frac{\omega_p^2}{(\omega^2 - \omega_c^2)}\left[\underline{\mathrm{E}} + \frac{i\omega_c}{\omega}\underline{\mathrm{E}} \times \hat{\underline{\mathrm{b}}} - \frac{\omega_c^2}{\omega^2}\left(\underline{\mathrm{E}} \cdot \hat{\underline{\mathrm{b}}}\right)\hat{\underline{\mathrm{b}}}\right] - \underline{\mathrm{E}} \tag{3.14}$$

This is simplest if expressed in a coordinate system with the z direction aligned to the static magnetic field $\hat{\underline{\mathrm{b}}} = \hat{\underline{\mathrm{z}}}$ and in terms of the refractive index $N = \frac{|k|c}{\omega}$ and unit vector of the propagation direction $\hat{\underline{\mathrm{n}}} = \underline{\mathrm{k}}/|k|$:

$$-N^2\left(\hat{\underline{\mathrm{n}}} \times \hat{\underline{\mathrm{n}}} \times \underline{\mathrm{E}}\right) \quad = \quad \frac{-\omega_p^2}{(\omega^2 - \omega_c^2)}\left[\underline{\mathrm{E}} + \frac{i\omega_c}{\omega}\left(E_y\hat{\underline{\mathrm{x}}} - E_x\hat{\underline{\mathrm{y}}}\right) - \frac{\omega_c^2}{\omega^2}E_z\hat{\underline{\mathrm{z}}}\right] + \underline{\mathrm{E}} \tag{3.15}$$

$$= \quad \left[\underline{\underline{1}} - \frac{X}{1 - Y^2}\begin{pmatrix} 1 & -iY & 0 \\ iY & 1 & 0 \\ 0 & 0 & 1 - Y^2 \end{pmatrix}\right]\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{3.16}$$

$$= \quad \underline{\underline{\epsilon}}' \cdot \underline{\mathrm{E}} \tag{3.17}$$

The electron plasma and cyclotron frequencies have been normalised to the wave frequency by $X = \omega_p^2/\omega^2$ and $Y = \omega_c/\omega$.

The matrix $\underline{\underline{\epsilon}}'$ is known as the dielectric tensor. The version derived here and shown in equation 3.16 is known as the cold plasma dielectric tensor and is valid only for low temperature plasmas due to the assumptions made in the use of the electron fluid momentum equation 3.8.

The solutions of equation 3.16 give the characteristic wave modes that can propagate in the given plasma. To assess the polarisation of these modes, it is more useful to work in a frame of reference with the z axis aligned to the propagation direction $\hat{\underline{\mathrm{k}}} = \hat{\underline{\mathrm{z}}}$ and magnetic field in the (x, z) plane $\underline{\mathrm{B}} = B\left(\sin\theta, 0, \cos\theta\right)$. This is done by an appropriate rotation of the dielectric tensor $\underline{\underline{\epsilon}}' \rightarrow \underline{\underline{\epsilon}}$:

$$\underline{\underline{\epsilon}} \quad = \quad \underline{\underline{1}} - \frac{X}{1 - Y^2}\begin{pmatrix} 1 - Y^2\sin^2\theta & -iY\cos\theta & -Y^2\cos\theta\sin\theta \\ iY\cos\theta & 1 & -iY\sin\theta \\ -Y^2 cos\theta sin\theta & iY\sin\theta & 1 - Y^2\cos^2\theta \end{pmatrix} \tag{3.18}$$

In this frame, equation 3.17 becomes:

$$\left(\underline{\hat{z}} \times \underline{\hat{z}} \times \underline{E}\right) N^2 - \underline{\underline{\epsilon}} \cdot \underline{E} = 0 \tag{3.19}$$

$$\left[\, N^2 \underline{\underline{1}} - N^2 \underline{\hat{z}}\, \underline{\hat{z}} - \underline{\underline{\epsilon}} \,\right] \cdot \underline{E} = 0 \tag{3.20}$$

Solutions exist where the determinant of the tensor is 0:

$$0 = \left| N^2 \underline{\underline{1}} - N^2 \underline{\hat{z}}\, \underline{\hat{z}} - \underline{\underline{\epsilon}} \right| \tag{3.21}$$

$$0 = N^4 - \left(\eta_{xx} + \eta_{yy}\right) N^2 + \eta_{xx}\eta_{yy} - \eta_{xy}\eta_{yx} \tag{3.22}$$

$$N^2 = \tfrac{1}{2}\left(\eta_{xx} + \eta_{yy}\right) \pm \tfrac{1}{2}\sqrt{\left(\eta_{xx} - \eta_{yy}\right)^2 + 4\eta_{xy}\eta_{yx}} \tag{3.23}$$

$$\text{where} \qquad \eta_{ij} = \epsilon_{ij} - \epsilon_{iz}\epsilon_{zi}/\epsilon_{zz} \tag{3.24}$$

Inserting the cold plasma dielectric tensor from equation 3.16 gives the refractive indices of the two characteristic modes:

$$N^2 = 1 - \frac{2X\left(1 - X\right)}{2\left(1 - X\right) - Y^2 \sin^2\theta \pm \sqrt{Y^4 \sin^4\theta + 4Y^2 \left(1 - X\right)^2 \cos^2\theta}} \tag{3.25}$$

3.2. Polarisation

3.2.1. Polarisation Description

The evolution of the components E_x and E_y of the electric field of a plane wave travelling in the $\hat{\underline{z}}$ direction is known as its polarisation. Several different representation can be used:

(Θ, Φ) - *Amplitude ratio and phase difference.*

As shown in figure 3.1a, the wave is considered as the sum of three waves with their E fields in each coordinate direction and the polarisation described by the amplitude ratio $\tan\Theta = |E_y|/|E_x|$ and phase difference Φ between the x and y components:

$$\begin{aligned} \underline{E} \;&=\; E_0 \cos\Theta \;\, e^{i(kz-\omega t)}\, \hat{\underline{x}} && (3.26) \\ &\quad +E_0 \sin\Theta \;\, e^{i(kz-\omega t+\Phi)}\, \hat{\underline{y}} \\ &\quad +E_z\,(z,t)\;\hat{\underline{z}} \\ & && (3.27) \end{aligned}$$

(ψ, χ) - *Principal angle and ellipticity angle.*

In general, the evolution of the vector (E_x, E_y) follows an elliptical trajectory as shown in figure 3.1b. The polarisation is described by the ellipticity angle χ and the angle between the major axis of the ellipse and the x-axis ψ.

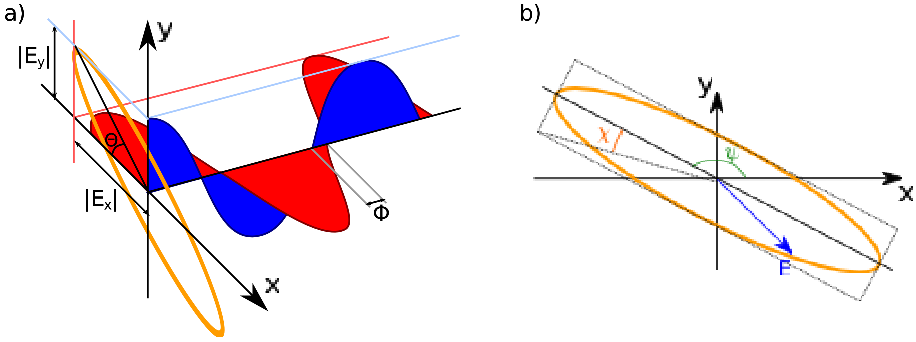

Figure 3.1.: Definition of wave polarisation in terms of a) Maximum amplitude ratio $\Theta = \tan^{-1}|E_y|/|E_x|$ and phase difference Φ between E_x and E_y, and b) ellipticity angle χ and principal polarisation angle Ψ.

$\underline{s}$ - *Stokes vector.*

The Stokes vector $\underline{s}$ can describe the polarisation with a 4 component vector in order to hold the polarisation state, radiation intensity and degree of polarisation. Here, waves are always considered fully polarised and total intensity is not desired so a 3 component vector with unit length is used. In terms of the electric field and the (ψ, χ) representation, this is:

$$\underline{s} = \begin{pmatrix} |E_x|^2 - |E_y|^2 \\ 2\Re\left(E_x E_y^*\right) \\ 2\Im\left(E_x E_y^*\right) \end{pmatrix} = \begin{pmatrix} \cos 2\chi \cos 2\psi \\ \cos 2\chi \sin 2\psi \\ \sin 2\chi \end{pmatrix} \tag{3.28}$$

Figure 3.2 shows a Stokes vector $\underline{s}$ and its relation to the angle and ellipticity on a unit sphere which is known as the Poincare sphere. Every point on the sphere represents a different polarisation and orthogonal polarisations exist at exactly opposite points.

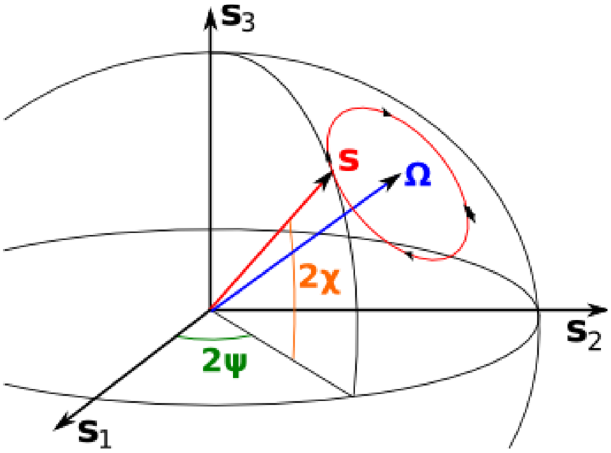

Figure 3.2.: The Stokes representation showing the polarisation $\underline{s}$ of an arbitrary wave on the Poincare sphere with polarisation angle ψ and ellipticity angle χ. Also shown is the polarisation evolution vector $\underline{\Omega}$ for an arbitrary plasma (aligned with one of the two characteristic wave modes) and the evolution of the wave $\underline{s}$ as it passes through that plasma.

The following relations allow the conversion of polarisation between the first two representations, and will be useful later:

$$\begin{aligned} \tan 2\psi &= \tan 2\Theta \ \cos\Phi & \qquad \cos 2\Theta &= \cos 2\chi \ \cos 2\psi \\ \sin 2\chi &= \sin 2\Theta \ \sin\Phi & \qquad \tan\Phi &= tan2\chi \,/\, \sin 2\psi \end{aligned} \tag{3.29}$$

3.2.2. Polarisation Evolution

The Stokes representation is particularly useful since it is known [53] that the polarisation $\underline{s}$ of a wave in a homogeneous medium evolves according to:

$$\frac{d\underline{s}}{dz} = \underline{\Omega} \times \underline{s} \tag{3.30}$$

where the *evolution vector* $\underline{\Omega}$ lies parallel to the Stokes vector of the fast characteristic wave $\underline{s}_1$ in the medium with magnitude proportional to the difference between the two refractive indices:

$$\underline{\Omega} = -\frac{\omega}{c}\left(N_2 - N_1\right)\underline{s}_1 \tag{3.31}$$

The Stokes vector for the propagating plane wave $\underline{s}$ follows a rotation around $\underline{\Omega}$ as it passes through the medium. Figure 3.2 shows the trajectory for an arbitrary homogeneous medium. For weakly inhomogeneous media, $\underline{\Omega}$ becomes $\underline{\Omega}(z)$ and equation 3.30 must be numerically integrated across slabs of thickness dz, giving $\underline{s}$ at the edge of each slab. To determine $\underline{\Omega}$ for a plasma, the Stokes vectors for the characteristic modes s_1 and s_2 must be found. To do this, it is useful to define the complex polarisation ratio p from the plane wave definition in equation 3.26, which can be expressed in the (Θ, Φ) description by:

$$p = \frac{E_y}{E_x} = \tan\Theta e^{i\Phi} \tag{3.32}$$

To determine p for each of the wave modes E_z must be eliminated from the wave equation 3.20, to give:

$$\begin{pmatrix} N^2 - \eta_{xx} & -\eta_{xy} \\ -\eta_{yx} & N^2 - \eta_{yy} \end{pmatrix} \begin{pmatrix} E_y \\ E_x \end{pmatrix} = 0 \tag{3.33}$$

Inserting the refractive indices of the two solutions (from equation 3.23), the polarisation ratio p is:

$$p = \frac{E_y}{E_x} = ig \pm i\sqrt{g^2+1} \tag{3.34}$$

$$\text{where } g = \frac{i\left(\eta_{xx} - \eta_{yy}\right)}{2\eta_{xy}} \tag{3.35}$$

From the cold plasma dielectric tensor (equation 3.16), the components η_{xy} and η_{yx} are imaginary and η_{xx} and η_{yy} are real. The implication is that g is real, p imaginary and that the phase difference between the x and y components is always $\pi/2$ for both of the characteristic modes:

$$\Phi = \tfrac{\pi}{2}, \qquad \tan\Theta = g \pm \sqrt{g^2+1} \tag{3.36}$$

Inserting the expression for $\underline{\mathrm{E}}$ (equation 3.26) with this Θ and Φ into the Stokes vector definition (equation 3.28) gives the Stokes vector for each mode:

$$\underline{\mathrm{s}}_{1,2} = \pm\frac{1}{\sqrt{g^2+1}}\begin{pmatrix} g \\ 0 \\ 1 \end{pmatrix} \tag{3.37}$$

Substituting this and the definition of g (equation 3.35) into the definition of Ω (equation 3.31) gives the evolution vector in terms of the dielectric tensor $\underline{\underline{\epsilon}}$:

$$\underline{\Omega} = \frac{\omega}{2c}\begin{bmatrix} \eta_{xx}-\eta_{yy} \\ 0 \\ -2i\eta_{xy} \end{bmatrix} \quad \text{recalling that} \quad \eta_{ij} = \epsilon_{ij} - \frac{\epsilon_{zj}\epsilon_{zj}}{\epsilon_{zz}} \tag{3.38}$$

Finally, inserting the cold plasma dielectric tensor gives $\underline{\Omega}$ in terms of the basic parameters n_e, $\underline{\mathrm{B}}$ and ω. However, the reference frame used so far has been with $\underline{\hat{\mathrm{x}}} \parallel \underline{\mathrm{B}}$ and since $\underline{\mathrm{B}}$ may arbitrarily change direction along the propagation path it is more useful to use a reference frame with x and y axes fixed relative to the lab frame. This might be based on the initial polarisation or on the equipment used to detect the final state. In such a frame, the magnetic field lies at an arbitrary angle $\underline{\mathrm{B}} = B_0\left(\sin\theta\cos\alpha,\ \sin\theta\sin\alpha,\ \cos\theta\right)$ and the final evolution vector is:

$$\underline{\Omega}^c = \frac{\omega_p^2\omega_c}{2c\omega^3\left(1-\omega_c^2/\omega^2\right)}\begin{bmatrix} \omega_c & \sin^2\theta\cos 2\alpha \\ \omega_c & \sin^2\theta\sin 2\alpha \\ 2\omega & \cos\theta \end{bmatrix} \tag{3.39}$$

3.2.3. Faraday Rotation and the Cotton-Mouton Effect

While the forward model is based on the integration of equation 3.30 and so is valid for any initial polarisation and any magnetic field with slow variation, it is useful to be aware of two specific cases:

$\underline{\mathrm{B}} \parallel \underline{\hat{\mathrm{z}}}$ - *Faraday Rotation*

With parallel magnetic field, $\theta = 0$ and Ω reduces to its third component. In this case the

Stokes vector of the propagating wave rotates around the $\underline{\hat{s}}_3$ axis as it progresses through the plasma. In terms of the polarisation ellipse (ψ, χ) this corresponds to a rotation of the initial polarisation state with constant ellipticity and is known as Faraday Rotation. The total rotation $\Delta\psi$ after passing through a plasma can be expressed approximately as a integral of the magnetic field magnitude and electron density along the path [53]:

$$\Delta\psi \approx \tfrac{1}{2} C_3 \int n_e(z)\, B_z(z)\, dz \tag{3.40}$$

$$\text{where } C_3 = \frac{q^3}{\omega^2 \epsilon_0 m^2 c} \tag{3.41}$$

$\underline{B} \perp \underline{\hat{z}}$ and $\psi_0 = 45^o$ - *The Cotton-Mouton Effect*

With a perpendicular magnetic field, $\theta = \pi/2$ and only the first two components of Ω may both be non-zero. The exact evolution depends heavily on the initial polarisation state but in the specific case where it is linearly polarised at 45^o to $\underline{B}$, the wave simply gains ellipticity. This is known as the Cotton-Mouton effect and the final induced ellipticity angle χ is given by [53]:

$$\chi \approx \tfrac{1}{2} C_1 \int n_e(z)\, B_\perp^2(z)\, dz \tag{3.42}$$

$$\text{where } C_1 = \frac{q^4}{2\omega^3 \epsilon_0 m^3 c} \tag{3.43}$$

In various cases, especially where either effect is weak or the polarisation change is small, it is possible to find the final polarisation approximately using equations 3.40 and 3.42 simultaneously, effectively treating the two effects as independent. Obviously, using this approach too far outside the assumptions used to derive the equations leads to inaccuracies in the final polarisation obtained. Unfortunately, this systematic inaccuracy is often viewed as an 'interference' or 'interaction' of the Faraday and Cotton-Mouton effects which should be removed by adding 'corrections' to the two equations. Such an approach leads to a confusing mix of models and terminology for situations where the concepts of Faraday rotation and Cotton-Mouton effect are not really relevant. Considerable effort has been spent evaluating which of these are more accurate and under what conditions they are valid [53][54][55][56][57] so that they may be used to convert measured polarisation directly into line integrals of the quantities of interest: n_e and B. The conceptual simplicity of the forward modelling approach is clear in this case, since it requires only equations 3.39 and 3.30 to recover all information possible about the physical quantities without loss of accuracy or generality.

3.3. The JET Interferometer/Polarimeter

The JET interferometer/polarimeter was originally developed as a Heterodyne Mach-Zander type interferometer used to obtain line integrated density along a number of lines of sight through the plasma[58]. The system has since been upgraded several times and now also provides polarisation measurements used to infer information about the magnetic field in the plasma[59].

Figure 3.3 shows a schematic of the system in its current form.

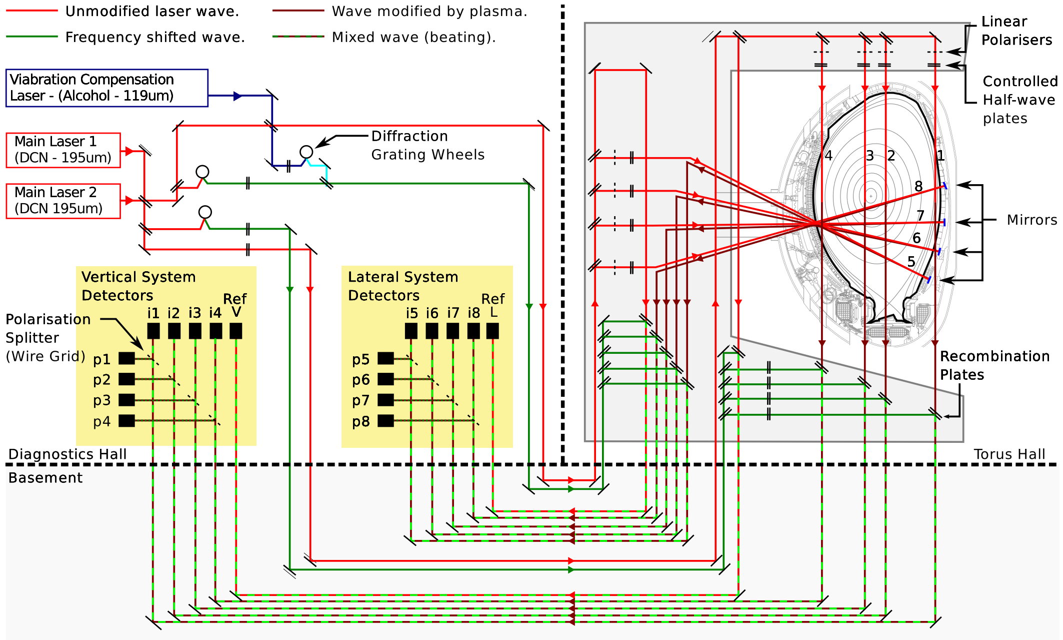

Figure 3.3.: Schematic of the JET combined Interferometry and Polarimetry system.

The system uses one of two 195μm Deuterium-Cyanide (DCN) lasers and a 119μm Alcohol laser. The DCN laser is the principal laser of the system and its beam is first split into two components, one designated for the system of vertical chords and the other for the lateral chords. For each system, the beam is split again and one component presented to a diffraction grating on the surface of a rotating wheel. The speed of the wheel is such that the motion causes a Doppler shift of the diffracted component by 100kHz. Both the shifted and pure beams pass through a series of mirrors to the Tokamak where the pure component is split into five, of which four then pass through the plasma. All four plasma *probing beams* and the fifth *reference beam* are then recombined with the 100kHz shifted beam. The interference of these gives a beat signal signal of approximately 100kHz. The five mixed beams return to the Diagnostics hall where they are split by a polarisation filter (a wire grid) into two orthogonal components which are each presented to an Indium Antimonide detector. The detectors measure a voltage proportional to the intensity of the ∼GHz wave, at sufficient temporal resolution to observe the ∼ 100kHz beat signal. For each pair of detectors, one remains from the original interferometer-only system and its signal is denoted $i(t)$ while the other was added for the polarimetry and is denoted $p(t)$. The four probing beams for the vertical measurements are denoted channels 1-4 and the reference beam *Ref-V* while four lateral plasma channels are 5-8 and the reference *Ref-L*.

The polarising wire-grid in front of the detectors is mounted at approximately 45^o to the beam line and reflects into the 'p' detectors the orthogonal component to the 'i' detectors. The coordinate system is defined so that 'i' detectors see the x component and the 'p' detectors see the y component.

3.3.1. The Heterodyne Interferometer

The wave arriving at the detectors is the sum of the shifted wave and either the plasma or reference waves. The electric field (x-component implied for the interferometer) of the frequency shifted (s), plasma (p) and reference (r) waves at a fixed z can be written as:

$$\begin{aligned} E_s &\propto e^{i((\omega+\Delta\omega)t+\phi_s)} \\ E_p &\propto e^{i(\omega t+\phi_p)} \\ E_r &\propto e^{i(\omega t+\phi_r)} \end{aligned} \tag{3.44}$$

Where ω is the laser frequency, $\Delta\omega$ is the angular frequency shift (2π 100kHz) and $\phi_{s,r,p}$ are the fixed or slow-varying phases on each signal so that ϕ_p includes the phase shift due to the plasma.

The detected signal is proportional to the amplitude of the complex field at each detector: $i(t) \propto EE^*$. Dropping the superscripts x, these are:

$$\begin{aligned} i(t) &= (E_p + E_s)(E_p^* + E_s^*) \propto 2|E_p||E_s|\cos(\Delta\omega t + \phi_p - \phi_s) + E_p^2 + E_s^2 \\ r(t) &= (E_r + E_s)(E_r^* + E_s^*) \propto 2|E_r||E_s|\cos(\Delta\omega t + \phi_r - \phi_s) + E_r^2 + E_s^2 \end{aligned} \tag{3.45}$$

A high-pass filter is used to remove the non-time varying components leaving only the ∼100kHz waves which are digitised and the phase difference between $i(t)$ and $r(t)$ extracted to give $\Delta\phi = \phi_p - \phi_r$. In a conventional (non-heterodyne) interferometer without the E_s component, the detected amplitude would vary with ϕ_p causing the signal to be entirely lost when the plasma density is such that the interference is entirely destructive. In this case, the 100kHz signal is always present at the detectors.

For each wave, the phase ϕ_p or ϕ_r at the detector is given by the integral of the refractive index along the complete path:

$$\Delta\phi = \frac{\omega}{c} \int N(z) dz \tag{3.46}$$

This will vary slowly in time (<< 100kHz) as the plasma refractive index changes, but also if the total integral length changes due to movement of the optical components in only one arm. The major source for the latter is from the in-vessel mirrors on the lateral channels. To isolate this motion, the entire system described is duplicated with the $119\mu m$ Alcohol laser and a frequency shift of 25kHz, using the same optics and the same detectors. The details are not covered here, but the two different frequencies allow the physical movement to be separated from the plasma change and then removed from the main system phase difference.

The phase difference due to the plasma can be obtained under the cold plasma approximation using the refractive index solutions in equation 3.25. For the JET system the expression can be simplified significantly: The propagation direction of all 8 channels lies almost entirely in the poloidal plane and the magnetic field is dominated by the toroidal field so the angle between them is large $\theta \approx \pi/2$ and terms in $\cos^2\theta$ small. Even for the high-field side vertical channel (1), the normalised cyclotron frequency is small $Y = \omega_c/\omega \sim 0.05$ and even for the core channels, the normalised plasma frequency is also small $X = \omega_p^2/\omega^2 \sim 0.003$. The first order term in both refractive index solutions is then simply $N \sim (1 - \frac{1}{2}X)$ and the phase difference becomes:

$$\begin{aligned} \Delta\phi &= \frac{\omega}{c} \int_0^{z_1} (1 - \frac{\omega_p^2}{2\omega^2}) dz \qquad (3.47) \\ &= \frac{\omega}{c} z_1 - \frac{q^2}{2c\epsilon_0 m\omega} \int_0^{z_1} n_e(z)\, dz \qquad (3.48) \end{aligned}$$

The first term is the normal path length and is simply absorbed into all the other constant phase shifts by defining $\Delta\phi = 0$ when the line integral density is zero at the start of the pulse. Unfortunately, because $\Delta\phi$ is periodic, if the plasma density increases enough the diagnostic will eventually present the same data as if it were zero. To mitigate this, the software assumes that the plasma changes only slowly and keeps track of how many complete periods or *fringes* it has passed. If the signal is lost or the density varies too rapidly, it can lose track of the fringe count causing a 'fringe jump'. The reported density integral will then be incorrect by an integer number of fringes, where a single fringe is:

$$\int_0^{z_1} n_e(z)\,dz = \frac{2c\epsilon_0 m\omega}{q^2}\,2\pi \tag{3.49}$$

$$= 1\,\mathit{fringe} = 1.143\times 10^{19} m^{-2} \tag{3.50}$$

While these events are common and are usually easily identified and corrected manually, no reliable automatic correction procedure has been developed so for any given pulse, some of the 8 channels of data may not be available. Where manual corrections have been performed, the line integrated density calculated from the phase difference (equation 3.47) is in practice very accurate.

Ideally, the entire system as described so far including the vibrational effects and with the full refractive index expressions should be used to model the likelihood function for the phase shift $P(\Delta\phi|n_e, \underline{\mathrm{B}}, ...)$. However, in almost all cases this will be calculated from, for example, a parametrisation of the spatial variation of n_e. The size and shape of the posterior $P(n_e|\Delta\phi)$ or anything calculated from it, will be dominated by the degeneracy in n_e due to having only effective line integrated measurements. In almost no cases will the size of the detailed likelihood (i.e the uncertainty of $\int n_e dz$ given $\Delta\phi$) be significant. For this reason, throughout this work, the line integral densities given by the standard analysis are taken as the 'data'. The appropriate line integral of the employed n_e parameterisation is the forward model and a Gaussian with the quoted standard uncertainty $\sigma_D \approx 10^{17} m^{-2}$ is taken for the likelihood function for each channel.

3.3.2. The Polarimeter - Model for the designed behaviour.

Given the definition of the (x, y) coordinate system, the components of the electric field of the plasma (p) and shifted (s) waves at the detectors are:

$$\begin{array}{ll} E_s^x \propto e^{i((\omega+\Delta\omega)t+\phi_s^x)} & E_p^x \propto e^{i(\omega t+\phi_p^x)} \\ E_s^y \propto e^{i((\omega+\Delta\omega)t+\phi_s^y)} & E_p^y \propto e^{i(\omega t+\phi_p^y)} \end{array} \tag{3.51}$$

Recalling the 'i' detector (x component) signal from equation 3.45 and stating the equivalent signal for the 'p' detector (y component), gives:

$$\begin{aligned} i(t) &= A_i|E_p^x||E_s^x|\cos\left(\Delta\omega t + \phi_p^x - \phi_s^x\right) + E_p^{x2} + E_s^{x2} \\ p(t) &= A_p|E_p^y||E_s^y|\cos\left(\Delta\omega t + \phi_p^y - \phi_s^y\right) + E_p^{y2} + E_s^{y2} \end{aligned} \tag{3.52}$$

where A_i and A_p are some fixed constant related to the detector sensitivity and gain etc.

These signals are high-pass filtered to remove the constant term and electronically multiplied and averaged over $\sim$25ms to produce the following signals:

$$\begin{aligned} RMS = \langle i \cdot i\rangle &\qquad PSD = \langle i \cdot p\rangle \\ RMP = \langle i' \cdot i'\rangle &\qquad PSP = \langle i' \cdot p\rangle \end{aligned}$$

where i' is produced by phase shifting the signal i by $\pi/2$, which also acquires some unknown gain $A_{i'}$.

In terms of the electric field components (but dropping the magnitude symbols), with the averaging over many cycles $t >> 2\pi/\Delta\omega$, these become approximately:

$$\begin{aligned} RMS &\propto A_i^2 E_p^{x2} E_s^{x2} \\ RMP &\propto A_{i'}^2 E_p^{x2} E_s^{x2} \\ PSD &\propto A_i A_p E_p^x E_s^x E_p^y E_s^y \cos\left(\phi_p^x + \phi_s^x - \phi_p^y - \phi_s^y\right) \\ PSP &\propto A_{i'} A_p E_p^x E_s^x E_p^y E_s^y \sin\left(\phi_p^x + \phi_s^x - \phi_p^y - \phi_s^y\right) \end{aligned}$$

These four signals are digitised and manipulated in software to produce the data signals R and R':

$$\begin{aligned} R &= \frac{PSD}{RMS} = \left(\frac{A_p E_s^y}{A_i E_s^x}\right)\frac{E_p^y}{E_p^x}\cos\left((\phi_p^x - \phi_p^y) + (\phi_s^x - \phi_s^y)\right) \\ R' &= \frac{PSP}{\sqrt{RMS\,.\,RMP}} = \left(\frac{A_p E_s^y}{A_i E_s^x}\right)\frac{E_p^y}{E_p^x}\sin\left((\phi_p^x - \phi_p^y) + (\phi_s^x - \phi_s^y)\right) \end{aligned}$$

The quantities of interest are those which describe the polarisation of the plasma beam $\tan\Theta_p = E_p^y/E_p^x$ and $\Phi_p = \phi_p^x - \phi_p^y$. All the remaining A, E and ϕ terms are unknown factors of the equipment which should be constant but are not known and so are collected into C and

ϕ_0:

$$\begin{aligned} R &= C^{-1}\ \tan\Theta_p\ \cos\left(\Phi_p+\phi_0\right) \\ R' &= C^{-1}\ \tan\Theta_p\ \sin\left(\Phi_p+\phi_0\right) \end{aligned} \tag{3.53}$$

Immediately before the input to the vessel, the plasma beams pass through a linear polariser and a half-wave plate. Since the input polarisation is linear, the phase difference at the detectors Φ_p must be zero when no plasma is present, regardless of the orientation of the detectors (and hence the coordinate system) with respect to the input polarisation. At the start of every JET pulse, before any plasma is introduced, the half-wave plates are swept through a large angle and then returned to their starting position. By fitting the observed R and R' signals to equation 3.53 for the sweep, it should be possible to determine the constants C and ϕ_0 along with the amplitude ratio (and hence polarisation angle) of the plasma beam, in the detector's (x,y) frame of reference.

3.3.3. The Polarimeter - Models for the observed behaviour

Unfortunately, the JET polarimeter system behaves in a way that is quite dissimilar to that described above. The primary indication of this is that the R and R' signals have a dependence on Θ_p during the calibration very different to that described by equation 3.53. Despite much investigation over several years, the cause of the anomalous behaviour has not been identified. To make some use of the data, while efforts to isolate the true problem continue, the theoretical model for the plasma is modified to match the empirically determined R, R' with some educated estimation of what the cause could be.

The simplest model which best fits most of the evidence (though not all) is the proposition that there is some birefringent optical element between the plasma and detectors[60] as illustrated in figure 3.4. Its optical axes are orientated at some unknown angle γ to the detector coordinate frame (with the detector frame at some unknown angle α to the plasma frame), and introduces an unknown phase shift ρ between those two components. Along with C, ϕ_0 and α of the system design, this description involves one too many unknown parameters than can be determined by the calibration sweep. Fortunately, it can be shown with some far from trivial trigonometry[61, 62, 63] that the single remaining degree of freedom in the parameters does not effect the relationship between the plasma beam polarisation Θ_p, Φ_p and the signals R, R'. This means the calibration sweep determines 4 parameters that can be used to determine Θ_p, Φ_p from R, R' although the values of this set of parameters do not reveal any details of the real physical set (C, ϕ_0, α, γ and ρ). Part of the JET polarimetry standard analysis code derives (see Appendix A) the plasma polarisation (given as ψ_p, χ_p) from the calibration fit using this procedure.

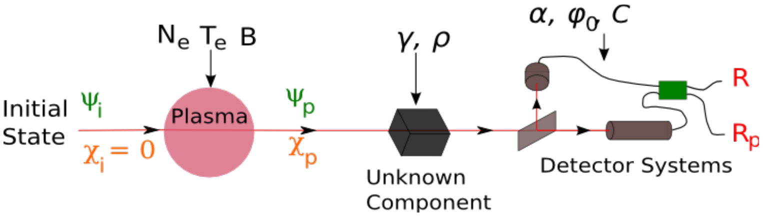

Figure 3.4.: Principle for 'unknown birefringent optics' model for the JET polarimetry diagnostic.

The 4 calibration factors are extremely well determined from the calibration sweep data. Taken with the very large signal to noise ratios of PSD, PSP, RMS, RMP, the uncertainty in $(\Theta_p, \Phi_p)/(\psi_p, \chi_p)$ should be much less than a degree on signals of up to 15 degrees and so the diagnostic should be extremely accurate, even with the unknown effects. Unfortunately, the model often can not accurately describe the observed calibration data and so can not be an accurate description of the system. It is this which causes the largest uncertainty in the calculated polarisation, something which is almost impossible to quantify directly.

With forward modelling and especially Bayesian analysis, this causes the problem that an accurate likelihood distribution cannot be constructed. Of course, this problem is just as real and as difficult to quantify for any non-Bayesian analysis procedure, though it is often less obvious that this is the case when using them. The most practical solution is to use a simple likelihood model, centred on the plasma polarisation as determined by the standard analysis (ψ_p, χ_p from section A). The width of this likelihood (the uncertainty) can be estimated by comparing this with the prediction of the plasma forward model from plasma parameters determined from other diagnostics. In this case, the plasma parameters to be determined are the electron density n_e and magnetic field $\underline{\mathrm{B}}$.

4. Interferometry And Polarimetry - Investigation and Analysis

4.1. Interferometry Results - Density Profile and Current Inversions

As discussed in section 3.3.1, the interferometry likelihood $P(\underline{\mathrm{D_i}} \,|\, n_e(R,Z))$ is sufficiently modelled as a fixed width Gaussian distribution, centred on the line integrated electron density along each line of sight. If the density at any spatial point $n_e(R,Z)$, is a linear function of the parameters used, the determination of those parameters from the (fringe-corrected) interferometry data $\underline{\mathrm{D_i}}$ satisfies the conditions of the truncated linear Gaussian inversion (TLGI - see section 2.3).

The 8 lines of sight available on JET are clearly insufficient to obtain any certainty on a complete 3D or even 2D (e.g. assuming axisymmetry) parametrisation. For low-flow equilibrium, electron density should be approximately constant on each closed flux surface (see section 1.3) which allows the density to be parametrised as a 1D function of normalised poloidal flux $n_e(\psi_N)$. For the SOL surfaces, the assumption may be slightly less valid, but since all but one of the lines of sight are very insensitive to the SOL density, the uncertainty will always be far larger than any systematic error due to this assumption.

4.1.1. Fixed equilibrium ψ_N with simple prior

To maintain linearity, the single fixed flux map $\psi_N(R,Z)$ taken from the standard equilibrium analysis (EFIT) is used and a simple linear interpolation of a series of knots at fixed ψ_N is used for $n_e(\psi_N)$. The values of n_e at the knots serve as the parameters. To allow the use of the TLGI, the prior must also be expressed as a Gaussian. The simplest form to use is an independent normal distribution on each knot, centred at zero with $+1\sigma$ slightly higher than JETs maximum operating regime and truncated so that $P(n_e < 0) = 0$ since densities cannot be negative.

Figure 4.1a shows the described inversion to $P\left(n_e \,|\, \underline{\mathrm{D_i}}, \psi_N\right)$ for a typical JET L-mode plasma, given the fixed flux surfaces of the standard magnetics-only equilibrium and the simple Gaussian prior.

These profiles look immediately 'incorrect' to any even mildly experienced Tokamak physicist, since it is usual to expect fairly smooth profiles throughout the plasma core, where there is no reason for cross-field transport or particle source rates to change significantly over short length scales. The posterior here does not in-fact exclude such profiles and shows that the short length scale information cannot be determined by the line integral measurements. Figure 4.1b illustrates this, showing that this degeneracy is correctly described by the posterior.

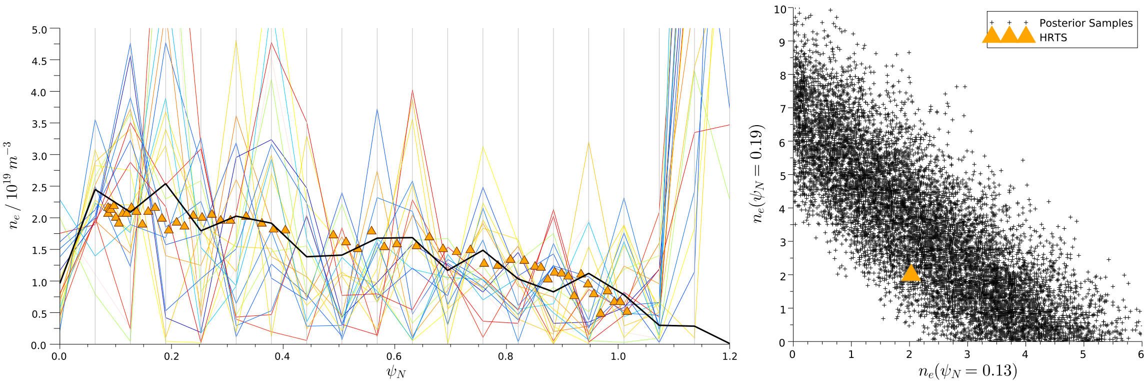

Figure 4.1.: a) Profiles at maximum (thick) and at 20 samples (thin) from posterior $P\left(n_e(\psi_N) \mid D_i,\ \psi_N^{EFIT}(R,Z)\right)$ using a simple uncorrelated Gaussian prior $P(n_e) = \mathcal{G}\left(n_e;\ 0, 10^{21}m^{-3}\right)$ for $n_e > 0$. b) Values of two neighbouring knots for 10000 samples showing degeneracy due to integrated measurements. The HRTS standard analysis results are shown with orange triangles.

The average n_e in the region of the two neighbouring nodes shown is fairly well determined, but the difference not. The posterior is correctly showing that with no other information, such large amplitude oscillations with short-wavelength could exist in the plasma.

4.1.2. Smoothing priors

Clearly, this claimed knowledge that transport and source rates do not change over short length scales should be included in the prior. To maintain applicability of the TLGI, this belief must be represented as a multivariate Gaussian over the parameters. A common approach is to include a Gaussian prior probability on the difference between each neighbouring pair of nodes. Each Gaussian is usually centred on 0, with a fixed width across the whole profile, or across separated regions such as the plasma core and edge. The effect is to smooth the profiles by constraining the first differential and hence it is known throughout this work as the *first differential smoothing prior*. The general expression is [1]:

$$\log P\left(y(x)\right) = -\frac{1}{2\sigma^2_{\frac{dy}{dx}}} \sum_{i=1}^{N-1} \left[\frac{(y_{i+1} - y_i)}{(x_{i+1} - x_i)} - 0 \right]^2 \tag{4.1}$$

An alternative is to apply a Gaussian to the change in this difference between each pair of

---

[1] As stated, equations 4.1 and 4.2 are *improper priors* because they do not integrate to a finite value. This would have no real consequence, since the likelihoods used here always ensure that the posterior is a proper PDF. However, in the software framework these priors are expressed as multivariate Gaussians and are required to be finite in order for it to manipulate the covariances. To do this, an uncorrelated normal distribution is added to the log P with σ set many orders of magnitude greater than the expected range in y, so the priors used are in fact, proper PDFs. Either way, this has almost no effect on the posterior.

nodes and the next, giving the *second differential smoothing prior*:

$$\log P\left(y(x)\right) = -\frac{1}{2\sigma^2_{\frac{d^2y}{dx^2}}} \sum_{i=2}^{N-1} \left[\frac{(y_{i+1} - y_i)}{(x_{i+1} - x_i)} - \frac{(y_i - y_{i-1})}{(x_i - x_{i-1})} - 0 \right]^2 \tag{4.2}$$

In this case, these can be applied either to the spatial derivatives or to the derivative with respect to the normalised flux.

The electron density gradient in steady-state is determined by the local electron particle transport and the local electron source rate and since there is no prior information about the magnitude of these, the first differential constraint, which prefers small gradients is hard to justify. However, the second differential constraint fits the belief that these gradients should not change rapidly over short scales. Since the parametrisation in this case is over normalised flux, it is most practical to constrain $d^2n_e/d\psi_N^2$. (The constraints are sufficiently weak that applying it to the true spatial differentials d^2n_e/dr^2 instead, has little effect on the results). Without a detailed model for the physics determining the transport and source rates, any exact value of $\sigma_{d^2n_e/d\psi_N^2}$ is hard to justify. In general, the value chosen is one that is strong enough to suppress wild variations in the profile due to diagnostic noise, but weak enough that the observed data remains inside the bulk of the likelihood distribution (i.e maintaining $\chi^2 \sim < 3$ for all the data). Usually, there is a range of one or two orders of magnitude that satisfies these conditions, and makes little difference to the inferred profiles.

Figure 4.2 shows the posterior obtained with the second differential smoothing prior with $\sigma_{d^2n_e/d\psi_N^2} = 1.0 \times 10^{21} m^{-3}$.

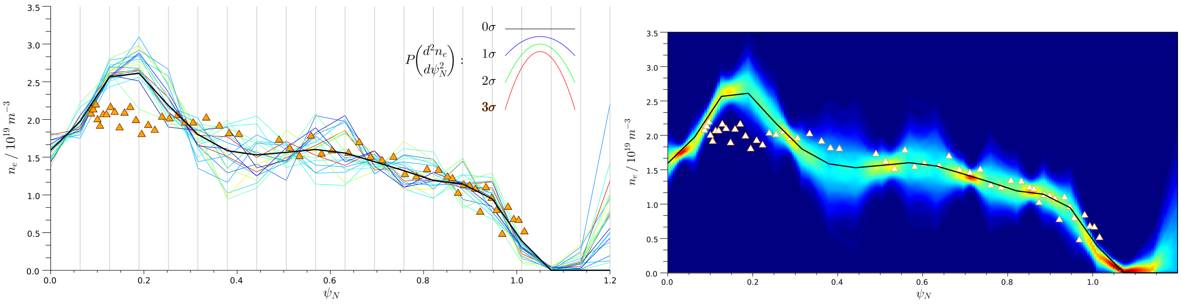

Figure 4.2.: a) Profiles at maximum (thick) and at 20 samples (thin) from posterior $P\left(n_e(\psi_N) \,|\, D_i,\, \psi_N^{EFIT}(R,Z)\right)$ using second differential smoothing prior. Also shown are curves at 1, 2 and 3σ of the prior used. b) Profile of marginal PDFs from the same posterior. The HRTS standard analysis results are shown with orange triangles.

Considering that only 8 data values relating to n_e have been used in the inversion and that the smoothing prior is relatively weak, the results of the inversion are surprisingly close to those of the HRTS diagnostic shown for comparison. For several values of ψ_N, the certainty is reasonably high, shown by the 'hot spots' in figure 4.2b. These correspond to the values of ψ_N for which one of the lines of sight is tangential to the surface, making that data value very sensitive to the density on that surface.

The remaining large discrepancy is only in the very core of the plasma ($\psi_N < 0.4$) where the posterior samples all show structure that is not seen in the HRTS profile, and which is difficult to see any cause of in this stable and quiescent L-mode plasma. Again, the posterior is in fact correct. It does show all reasonably probable n_e profiles, given ψ_N. If the profiles are considered improbable, the implication is that the fixed ψ_N is improbable (or possibly one of the other assumptions). The large uncertainty for $\psi_N > 1.15$ appears because the density here has little effect on any of the channels. Some channels enter this region only briefly while others pass behind the first wall before this level of ψ_N is reached and so the model does not include them at all. Under the assumption that there will be no significant electron density in the toroidal shadow of the first wall, in the interferometry beam passages, or on flux surfaces that are only inside the first wall for some channels, a prior can be included that constrains this region to very low or zero density. However, given that channel 1 already constrains $n_e(\psi_N \approx 1.1)$ to very low density, this has little effect on the rest of the profile.

4.1.3. Interferometry and current tomography - Free ψ_N non-linear inversion.

If the electron density is the object of interest, then the posterior really desired is $P\left(n_e \mid D_i\right)$ which is equal to that found in the last section $P\left(n_e \mid \underline{\mathrm{D_i}},\ \psi_N\right)$, integrated over all possible flux surface arrangements. This would be as uninformative as assuming nothing of the 2D geometry and parameterising $n_e(R, Z)$ directly and so the PDF would give almost any possible electron density. However, some information about the flux surface geometry can be gained from the magnetic diagnostics by including their data D_m to give the posterior $P\left(n_e \mid D_i,\ D_m\right)$. Practically, this is a combination of the current tomography (CT) from section 2.5 and the density inversion from the last section with ψ_N calculated from the toroidal current beams $\underline{\mathrm{j}_\phi}$. The toroidal currents $\underline{\mathrm{j}_\phi}$ and electron density knots $\underline{\mathrm{n_e}}$ form the complete set of parameters.

This unfortunately introduces the non-linear relationship between the interferometry data D_i and the currents j_ϕ, since a change to j_ϕ changes both the normalisation and geometry of ψ_N. This non-linearity turns out to be weak enough that the LGI procedure can still be performed and while the resulting Gaussian is not a good enough approximation to the real PDF to draw samples directly, it does give the approximate global shape, especially for the strongest correlations. After appropriate rescaling, it performs excellently as a proposal distribution for the Metropolis Hastings MCMC (see section 2.4.5), something that is otherwise extremely hard to obtain in the high dimensionality necessary to correctly describe the toroidal currents j_ϕ. Practically, the use of the approximate TLGI reduces the execution time of the MCMC by over an order of magnitude.

Figure 4.3 shows the density profiles from the full posterior. The inversion is to a parametrisation of 50 knots of $n_e(\psi_N)$, 217 plasma current beams spread inside the first wall and 48 Iron-core current parameters. The priors used were the second differential prior on n_e with $\sigma_{d^2 n_e/d\psi_N^2} = 1.0 \times 10^{21}$ and a very weak Conditional-Auto-Regressive (CAR) prior over the plasma currents.

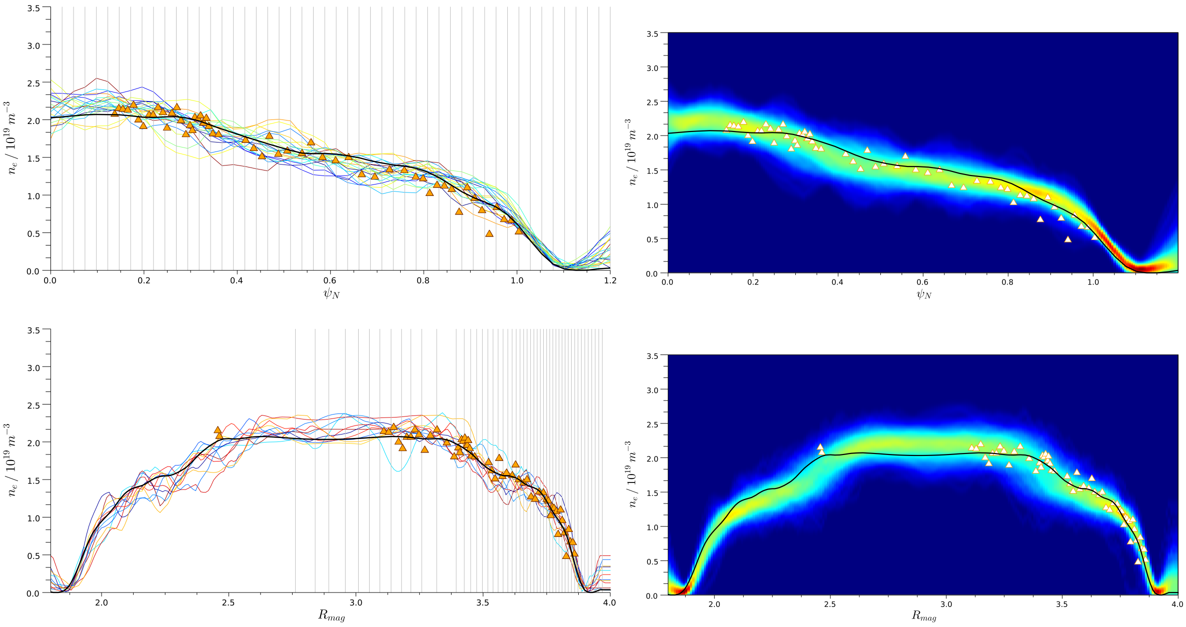

Figure 4.3.: a/c) n_e Profiles for 20 samples and b/d) profile marginals from posterior $P\left(n_e(\psi_N), j_\phi \mid D_i,\, D_m\right)$ using second differential smoothing prior and free plasma currents, versus a/b) ψ_N and b/d) R_{mag}. The HRTS standard analysis results are shown with orange triangles.

As well as the density profile, each sample of $P\left(n_e,\, j_\phi \mid D_i,\, D_m\right)$ contains a complete description of the toroidal current. Comparing this to the posterior from current tomography alone $P\left(j_\phi \mid D_m\right)$, indicates the information that the interferometry data (and the n_e prior) provides about the currents and hence the magnetic geometry. Figure 4.4 shows the separatrix and magnetic axis position from samples of the two posteriors to demonstrate the principle. In this case, the apparent accuracy might be easily obtained simply by increasing the prior assumptions (e.g. using a strong CAR prior, or assuming equilibrium) but the improvement here has been derived directly from the data of a diagnostic not normally considered capable of measuring anything about the plasma current.

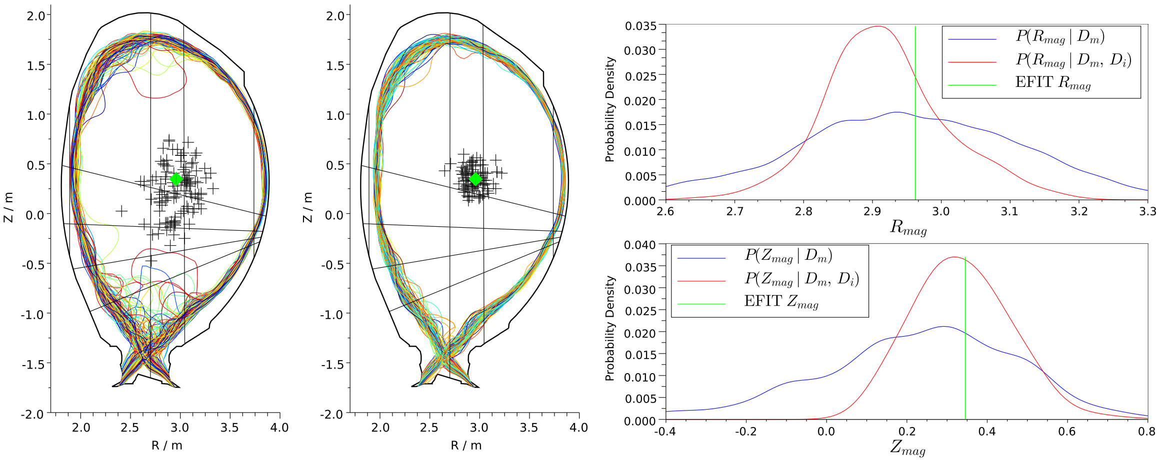

Figure 4.4.: Comparision of magnetic configurations of the posterior from Current Tomography (CT) alone and including interferometry data and electron density profile prior, with the same parameterisation and prior for toroidal currents. a) Separatrix and magnetic axis position for 100 samples from posterior for left) CT only and right) CT+interferometry. Interferometry lines of sight and first wall (black) and standard magnetics EFIT axis position (green diamond) are also shown. b) PDFs for top) R_{mag} and bottom) Z_{mag} each showing CT-only (blue), CT+interferometry (red) and EFIT (green).

At first glance the results are concerning, since it is hard to believe that 8 numbers can provide information about the plasma currents at the same time as electron density profiles that appear to have a similar resolution as the HRTS system. It should be emphasised though, that the electron density priors (the smoothing and assumption of constancy on flux surfaces) are providing much of the information and the results are accurate because the priors are good priors for L-mode plasmas.

For H-Mode plasmas, the transport rates change rapidly across the plasma edge and hence the smoothing prior must be relaxed, even if only in this region. Unfortunately, JET H-mode shots are usually positioned in such a way that the interferometry line of sight which measures only SOL density (channel 1) is further from the plasma edge, as seen in figure 4.5b. The pedestal information is effectively provided by this channel and the difference between the two core channels (channels 2 and 3) which couples the density near the X-point to the plasma core. To illustrate this, figure 4.5a shows the integration weights and density profile (at the posterior maxima) for the L-mode plasma of figure 4.3 and for a typical H-mode plasma. It is clear that the pedestal region of the H-mode pulse is illuminated by even less information than in the L-mode case. The only information about the profiles between around $\psi_N = 1.0$ and $\psi_N = 1.2$ is in the prior and so almost all samples drawn from the PDF have a constant gradient in that region. This has a knock-on effect on the core, due to the differences between channels 2 and 3, which can be seen by comparing the posterior profile marginals to the HRTS standard analysis in figure 4.6.

Clearly, it will not be possible to infer the shape of the H-mode pedestal from the interferometry alone. However, the data does hold some information about it and this subtle information content is relied upon in chapter 5, in which the interferometry model is used with forward models developed for a pair of diagnostics which, taken independently, provide shape detail with unknown absolute magnitude.

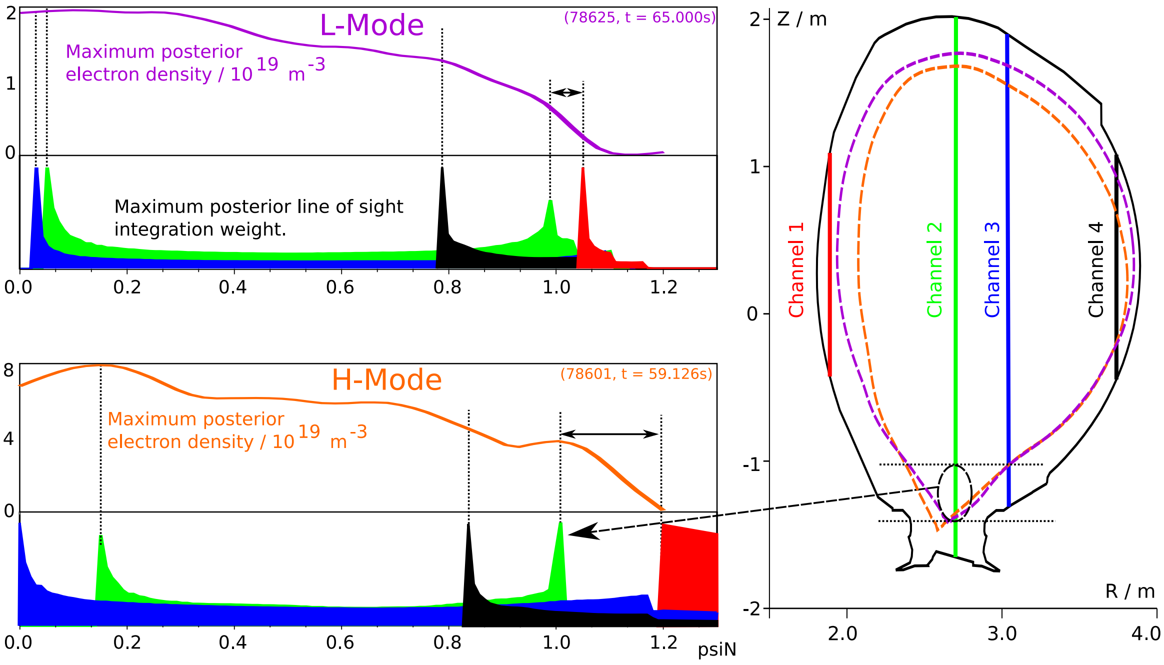

Figure 4.5.: a) Integration weights over ψ_N and $n_e(\psi_N)$ at posterior maximum for typical L-mode (top) and H-Mode (bottom) plasmas. b) Separatrix ($\psi_N = 1$) for both plasmas relative to lines of sight.

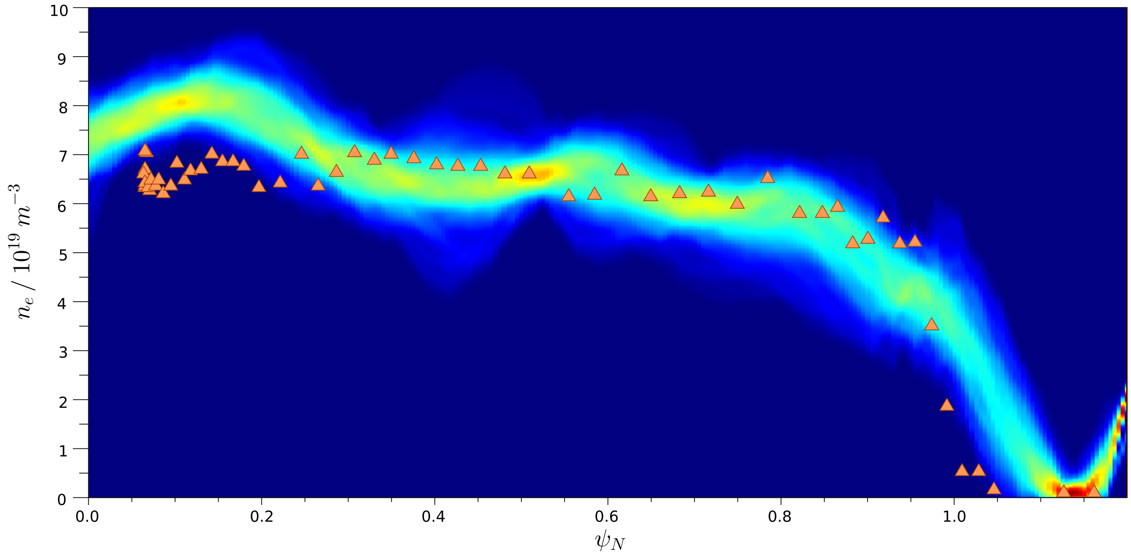

Figure 4.6.: Profile marginals from posterior $P\left(n_e(\psi_N), J \mid D_i,\ D_m\right)$ for H-Mode plasma using second differential constraint 'smoothing' prior and free plasma currents.

4.2. Accuracy of the Polarimetry Diagnostic Model

As discussed in section 3.3.3, the polarimetry model is split into two stages: the plasma model which gives the polarisation of the wave leaving the plasma (ψ, χ) and JET system model which determines the data (R, R'). The former is assumed to be accurate here (at least for low T_e plasmas) as it is used regularly in plasma polarimetry on other Tokamaks and plasma experiments. The diagnostic model is an ad-hoc model, constructed to fit behaviour that is not understood. In order to examine the validity of this part, this section compares the measured polarimetry data with predictions of the model, based on electron density and magnetic field information obtained from the interferometry and magnetics.

4.2.1. Statistical Comparison

The full two-stage forward model could be used to predict the data (R, R') and this compared to the measured (R, R'). Instead, the first stage is used to determine the predicted polarisation after the plasma (ψ_p, χ_p) and this is compared to (ψ_d, χ_d) - the polarisation after the plasma derived from the measured data (R, R') using the diagnostic system model (by the standard analysis code). The comparison is made in (ψ, χ) space because they retain some physical relevance to the plasma. In the first approximation, ψ_p is described by Faraday rotation (equation 3.40) which is sensitive to n_e and the beam-parallel field $B_{\|}$ which is the poloidal field, given the channel set up. The ellipticity χ_p is approximated by the Cotton-Mouton effect (equation 3.42) which is sensitive to n_e and $B_{\perp}$. $B_{\perp}$ is dominated by the the accurately known vacuum toroidal field (the field created by the toroidal field coils) which for the vertical channels is constant as it varies only with major radius. The consequence is that, for the vertical channels, χ_p is proportional to line-integrated n_e and so should be well determined by the interferometry and not strongly effected by less accurately known poloidal magnetic field.

Ideally, the forward model would be used on samples of the full non-linear CT + Interferometry inversion of 4.1 so that the uncertainty in the prediction includes both the uncertainty of n_e and of B given both the interferometry and magnetics. Unfortunately the non-linear algorithms are too computationally expensive to allow a good statistical comparison on a large number of time points. Instead, ψ_N and B are fixed to the EFITJ solution so the prediction uncertainty includes only that on n_e from the fixed ψ_N interferometry inversion. The systematic error this was shown to introduce only significantly affects ψ_N and B in the plasma core (which is seen qualitatively in the results here and proven later when analysing the equilibrium uncertainty itself with Bayesian methods). The distribution of predictions produced is therefore $P(\chi_p, \psi_p \mid D_i, J^{EFIT})$.

Figure 4.7 shows the prediction/measurement comparison at 14ms intervals over a typical H-Mode JET pulse and the statistical comparison including 50000 time points over $\sim$ 1000 pulses covering several years of JET data.

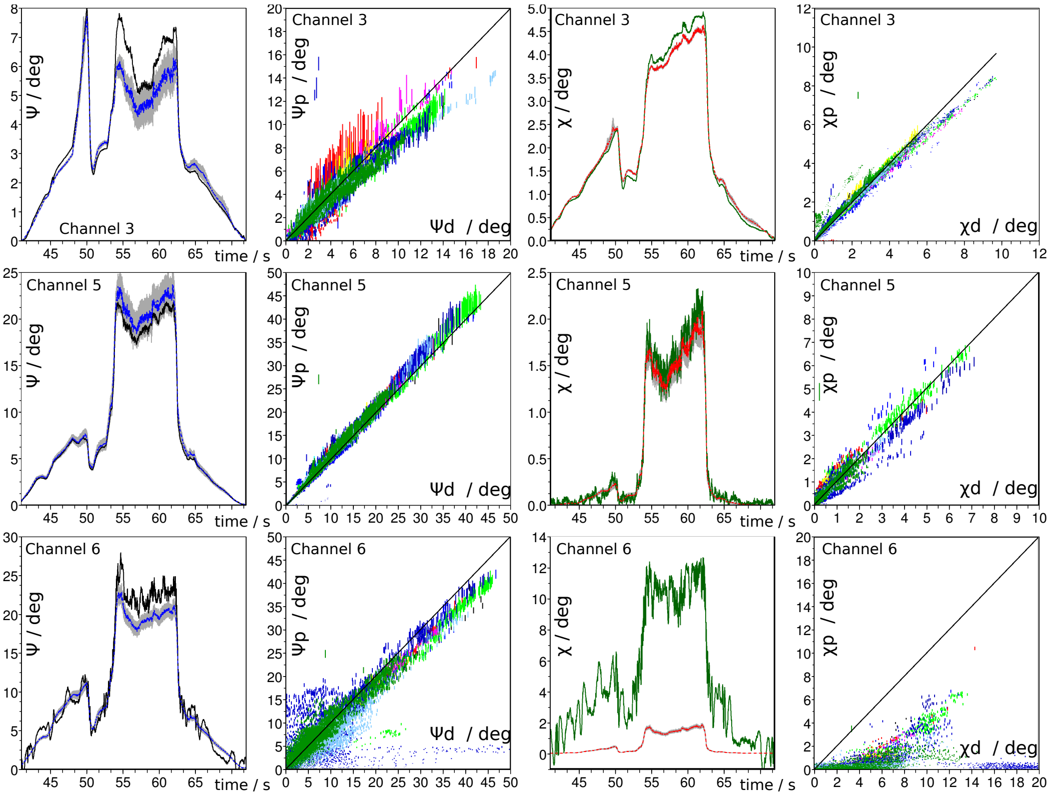

Figure 4.7.: : Column 1: Principal polarisation angle ψ after plasma comparing prediction from EFIT and Interferometry mean (blue) and uncertainty (gray) with the polarisation angle derived from data (black) for JET pulse 70545. Column 2: Predicted versus data derived ψ for $\sim$ 1000 pulses. Column 3: Prediction mean (red) and uncertainty (gray) of ellipticity angle χ compared to data derived value for the single pulse. Column 4: Predicted versus data derived χ for $\sim$ 1000 pulses. Rows show channel 3 - core channel sensitive to uncertainty in EFIT solution, channel 5 - Example of good calibration fit and channel 6 - failed calibration fit. Points in scatter plots are coloured according to experimental campaign.

It is immediately clear that the that the data derived polarisation (ψ_d, χ_d) is far outside the uncertainty in the prediction (ψ_p, χ_p) in many cases. For ψ, the general trends appear to be correct and the statistics show approximate agreement with some spread. For χ, the situation is far worse, with only channel 5 showing agreement within uncertainty and channel 6 showing none at all. It is possible to believe that the disagreement for ψ of channel 3 is entirely due to inaccuracy in the equilibrium solution as it passes close the core and the disagreement exists only during H-mode $(54s < t < 62s)$ where EFIT is unable to properly describe the current and pressure profiles. However, this cannot be argued for χ which is not sensitive to the poloidal field or flux geometry on any channel. It also does not explain ψ on channel 6, which does not pass near the core. Channel 5 is the single good case, as it shows largely consistent agreement in both ψ and χ.

Channel 6 of figure 4.7 also shows a large oscillation-like variation in both ψ_d and χ_d. Plotting either of these, or the raw signals, against the line integrated density shows that there are clear sinusoidal oscillations in $\int n_e dl$, with a wavelength of almost exactly one fringe. The appearance of these can be derived, by adding the term $\delta e^{i\omega t}$ to E_s in equation 3.44, representing a non-shifted component with small amplitude δ, in the signal that is supposed to be frequency shifted. This shows that it would interact with the plasma channels to give a variation with density, behaving similar to a simple non-heterodyne interferometer. Whether or not this is the real cause of these oscillations is not yet known but it does predict that they vary slightly in phase and amplitude with both ψ_d and χ_d, which is also observed in the data. This makes their isolation and removal by post-processing difficult.

Given that the mapping $(R, R') \rightarrow (\psi, \chi)$ is a complex combined operation based on the unknown optics model and the diagnostic calibration, the severe disagreement and oscillations seen in χ should also raise suspicion of ψ on those channels and pulses. Any useful attempt to analyse the JET polarimeter data must consider the effects of the diagnostic model itself. Despite this, many attempts[54, 55, 57] have been made to asses the accuracy of different approximations of the plasma polarisation evolution by comparing them with the data-derived (ψ_D, χ_d), without any consideration of the diagnostic's behaviour or calibration.

For completeness, figure 4.12 (page 85) shows the traces and statistics for the remaining channels.

4.2.2. Examination of Calibration Mapping

The right pane of figure 4.8 shows the measured signal (R, R') for the pulse shown in figure 4.7 and a regular grid in that space. The left pane shows the same signal and grid mapped to (ψ, χ) space according to the diagnostic model and calibration, compared with the mean prediction trace. The full signal includes the calibration sweep of input polarisation angle before the plasma starts. This sweep is in ψ only with $\chi = 0$ and the calibration procedure tries to fit the parameters so that the data calibration trace (magenta) maps to the $\chi = 0$ line (cyan).

For channel 5, the procedure clearly works well. The calibration sweep has been reasonably well fitted to the real calibration sweep line and the plasma part of the trace (red) matches the prediction (blue) well, explaining the agreement seen in figure 4.7. For channel 6, shown

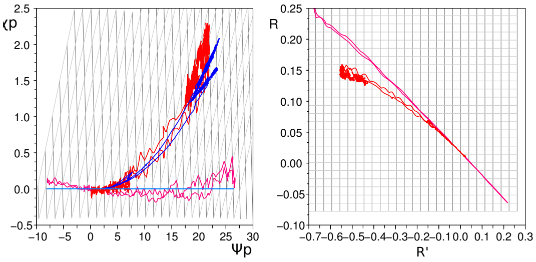

Figure 4.8.: Channel 5: Good calibration.

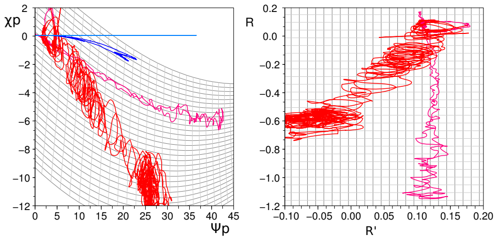

Figure 4.9.: Channel 6 - Poor calibration.

Each shows left) the mean predicted polarisation during plasma (blue) and during calibration sweep (cyan) compared to mapped data during plasma (red) and during calibration (magenta). right) Measured data in (R, R'). Both graphs also show a regular (R, R') grid.

in figure 4.9, the calibration has almost entirely failed. The sweep signal has not been fitted to the real sweep and so it is not surprising that the plasma trace does not agree either. The large oscillation is also clearly present in the measured data.

Figure 4.10 shows the mapping for channel 3. In this case, the calibration signal has been successfully matched to the real sweep and the Ohmic part of the plasma ($\chi \sim> -3$) appears to be matched well. If the disagreement in the H-Mode part of the pulse is to be blamed on the calibration, a very localised distortion to the calibration mapping would be required and so it seems likely that the mapping is in fact reasonably accurate and that it is the inaccuracy of the EFITJ solution in the plasma core which causes the disagreement in this case. It is not clear however, how accurate the ψ_d derived by the calibration is, given that it is so far from the $\chi = 0$ calibration line.

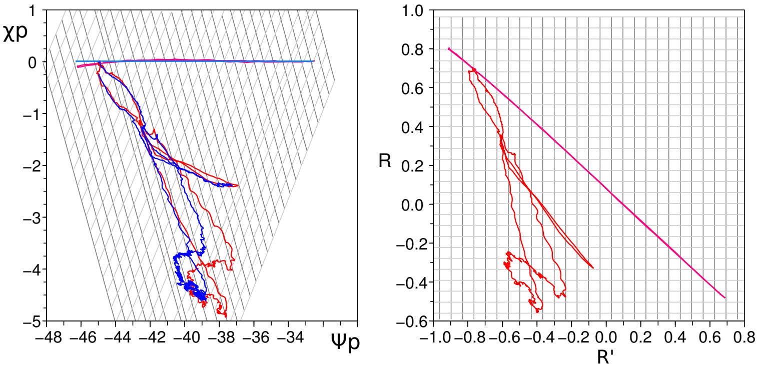

Figure 4.10.: Mapping as in figure 4.8, but for channel 3 passing through the plasma core.

Unfortunately, until the unknown effects in diagnostic which cause the errors in the $(\psi, \chi) \longleftrightarrow (R, R')$ mapping are resolved, or a more rigorous model is developed, the uncertainty on (ψ_d, χ_d) has to be estimated in each case by studying the mapping in this way.

4.2.3. Predictions from Current Tomography

The Current Tomography method in principle provides a way to determine if the disagreement in channel 3 during H-Mode could be due to inaccuracy in j_ϕ, since it provides j_ϕ with uncertainty. Ideally, the data-derived ψ_d would be compared to a prediction from the full posterior $P(n_e, j_\phi \mid D_m, D_i)$ but, as mentioned earlier this is not practical for a large number of time points. However, it is possible to find an over-pessimistic (larger uncertainty) approximation by drawing samples from the magnetics only inversion $P(j_\phi \mid D_m)$ including the CAR prior and then performing the fixed ψ_N interferometry inversion to $P(n_e \mid D_i, j_\phi)$ on each sample of the first. Both inversions can be performed using the (T)LGI and so drawing samples is inexpensive. The PDF described by the samples (which does not relate to any rigorously definable one) is conceptually similar to $P(n_e, j_\phi \,|\, D_m, D_i)$ except in that it does not include the information that the interferometry and n_e priors give about j_ϕ.

Figure 4.11 shows the mean and $\pm 2\sigma$ of ψ_p from these samples and the data-derived ψ_d for an H-Mode pulse compared to the same from EFIT for channels 3 and 5.

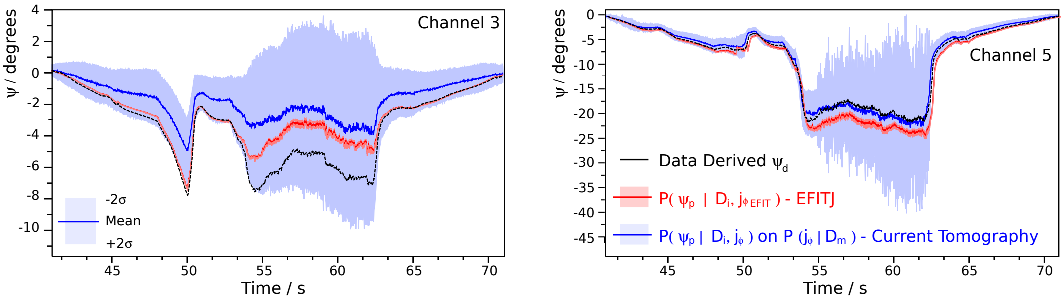

Figure 4.11.: Data-derived principal polarisation angle ψ_d (black) compared to predictions ψ_p from EFITJ solution (red) and Current Tomography $P(j_\phi \mid D_m), P(n_e \mid D_i, j_\phi)$ (blue).

The spread of ψ_p predictions does include the data-derived ψ_d for the H-Mode section, showing that ψ_d is not incompatible with the magnetics and interferometry data under the CT model and CAR prior. However, the large spread covers almost any sensible value and far exceeds the disagreement between EFITJ and the polarimetry data. While this does mean that the CT and CAR prior do not help quantify the accuracy of the polarimetry calibration, it suggests that if the accuracy were known, including the polarimetry data in the CT system would provide a significant amount of extra information about j_ϕ, since it will exclude a large part of $P(j_\phi \mid D_m)$.

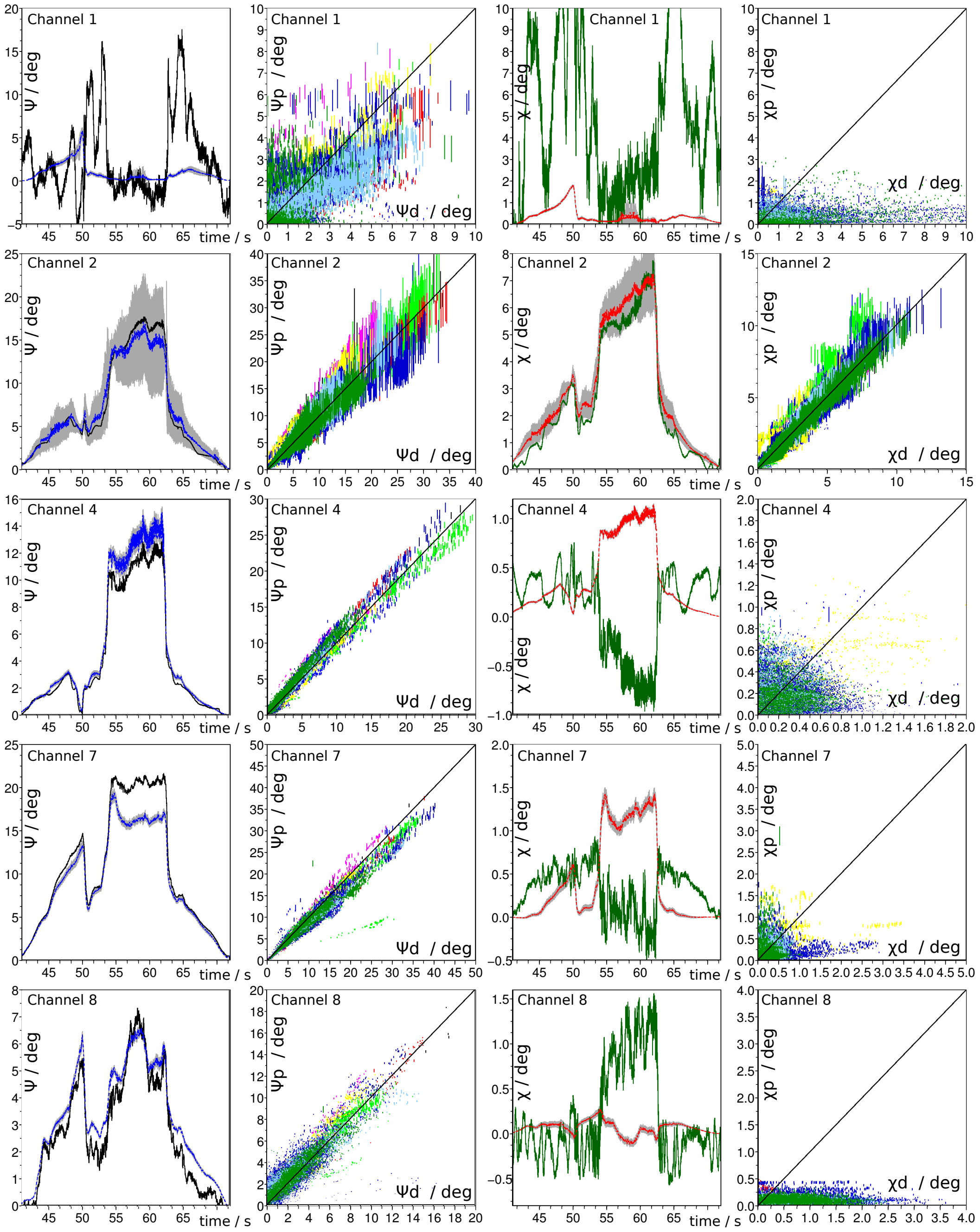

Figure 4.12.: Trace and statistical comparisons as in figure 4.7 for channels 1, 2, 4, 7, 8

4.3. Comparison of High Temperature Models

4.3.1. Finite Temperature Effects

The plasma polarisation evolution model used so far was based on the electron fluid momentum equation (equation 3.8) with the pressure and collisional terms neglected, which is the cold plasma approximation. In more thorough treatments (see for example [52]) which begin from the full momentum equation and assume a Maxwellian (relativistic or classical) electron velocity distribution, the dielectric tensor is found to have terms including the temperature T_e, these are known as *finite temperature effects.*

Recently, the effect of finite electron temperature T_e on the polarisation evolution has been studied, giving 'warm plasma' approximations to first order in normalised electron temperature $\tau = T_e/(m_e c^2)$ in the non-relativistic[64] and weakly relativistic[65] limits. While these approximations had previously been used in the evaluation of plasma polarimetry[54], they had never been verified by experimental observation. This section compares the data derived values of induced ellipticity χ_d, taken from the JET Polarimeter standard analysis, with predictions of the forward model based on the cold plasma as in the last section and using the non-relativistic and relativistic approximations given by each of these papers.

For high T_e plasmas, electrons with high thermal velocity experience a Doppler shifted incident wave and so their contribution to the plasma's effect on the wave is modified. This is addressed in [64], where the evolution vector $\underline{\Omega}$ is calculated from the non-relativistic plasma dielectric tensor. An asymptotic expansion is used for the plasma dispersion function and terms above first order in τ are dropped. The result, written in terms of the cold plasma approximation $\underline{\Omega}^c$, is the non-relativistic warm plasma approximation $\underline{\Omega}^n$:

$$\underline{\Omega}^n \approx \underline{\Omega}^c + \tau \begin{bmatrix} 12 & \Omega_1^c \\ 12 & \Omega_2^c \\ 3 & \Omega_3^c \end{bmatrix} \tag{4.3}$$

It is argued in [64], that the non-relativistic limit is sufficient to describe plasmas up to at least $T_e = 15$keV. However, it was later shown in [65] that the relativistic mass increase of high velocity electrons affects their response to the wave's electric field by a similar magnitude to the non-relativistic finite T_e effects. A calculation including both effects to first order in τ is given using a new iterative technique. Written in terms of the cold approximation $\underline{\Omega}^c$ and assuming the refractive index $N \approx 1$, this relativistic warm approximation $\underline{\Omega}^r$ is given[65] as:

$$\underline{\Omega}^{r} \approx \underline{\Omega}^{c} + \tau \begin{bmatrix} 9/2 & \Omega_1^c \\ 9/2 & \Omega_2^c \\ -2 & \Omega_3^c \end{bmatrix} \tag{4.4}$$

4.3.2. Model Evaluation

As before, the forward model is evaluated from n_e profile samples taken from the posterior $P(n_e \mid D_i, j_\phi^{EFIT})$. T_e is provided by the core LIDAR TS standard analysis code. These T_e measurements are prone to large random fluctuation on individual data points, occasionally giving unphysical large values which result in large jumps in the temperature effects calculated from 4.3 and 4.4. As the T_e values are provided with a well calculated uncertainty, the T_e profile as a function of ψ_N is modelled as 3-point spline function and fit to the LIDAR data points. Given that the theoretical T_e effects are relatively small, and so T_e is required to an accuracy of only around 1keV, any systematic effect this procedure causes will be negligible.

For analysis of the high T_e effects, channel 3 is used as it has a reasonably stable calibration and passes through the plasma core where T_e is typically largest (up to approximately 12keV). For the full range of JET plasmas, this channel measures rotations of $\Delta\psi < 15^o$ and ellipticities of $\chi < 8^o$. The beam-perpendicular magnetic field $B_\perp$ is dominated by the toroidal field and the channel is set up with initial polarisation at 45^o to this. The channel therefore approximately satisfies the assumptions of equations 3.40 and 3.42. Further to this, $B_\perp$ is dominated by the externally applied 'vacuum' field which varies with major radius as $1/R$ and since channel 3 lies at constant R, the $B_\perp$ term is effectively constant and can be taken outside the integral. This means χ is approximately proportional to line integrated density, which is known to an accuracy of less then 1% from the interferometry system. It should therefore be possible to calculate χ to a greater accuracy than the difference between the three models.

As shown in section 4.2, the rotation $\Delta\psi$ is heavily dependent on the poloidal magnetic field which is determined from the equilibrium code. As this is based only on magnetic measurements far from the centre of the plasma, it is heavily determined by the constraints made on the pressure and current profiles which can result large inaccuracies. The primary effect on channel 3 of any inaccuracy in the equilibrium reconstruction is effectively just the incorrect positioning of the magnetic axis. It is therefore possible to approximately include it in the ψ/χ uncertainty by varying the line of sight position in the poloidal plane randomly by 4cm about its prescribed position for each n_e sample.

For χ, the calculation uncertainty determined in this way is typically 0.4%, which is signifi-

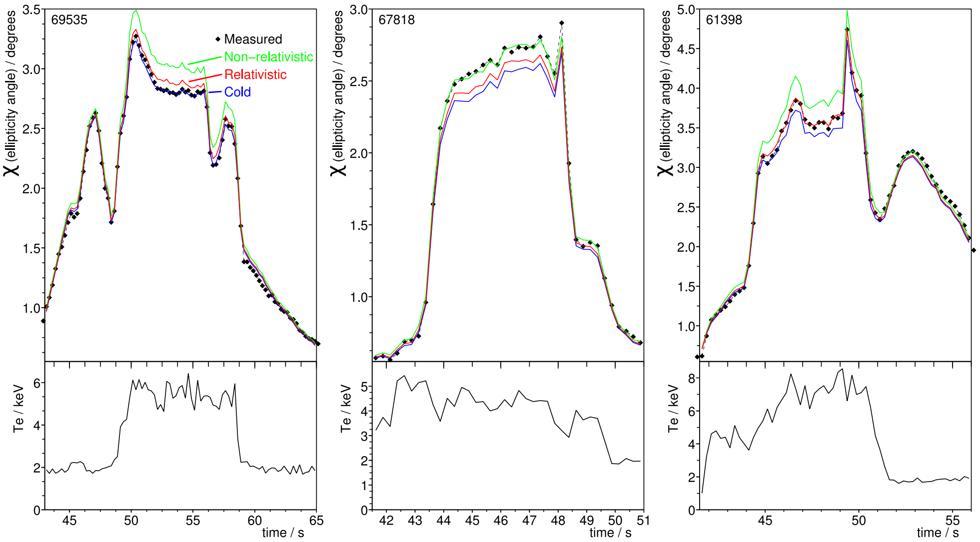

Figure 4.13.: Time traces for three high T_e JET pulses showing data-derived induced ellipticity χ_d (points) and ellipticity calculated from the the cold plasma χ^c (blue/lower line), non-relativistic χ^n (green/top line) and relativistic χ^r (red/middle line) approximations. The three pulses show apparent agreement with a) χ^c, b) χ^n and c) χ^r. Measurement uncertainty is principally from calibration so not shown and prediction uncertainty due to n_e, T_e, B is $\sigma_\chi \approx 0.4\%$ for all points (not shown for clarity).

cantly smaller than the associated temperature effect of approximately 2% for 10keV. For $\Delta\psi$, the typical calculation uncertainty is 10% which is much larger than the typical temperature effects of less than 1% at 10 keV. Therefore, in this section, only the induced ellipticity χ is studied.

4.3.3. Pulse selection

Unfortunately, because the error on the data-derived plasma induced values (ψ_d, χ_d) comes primarily from the calibration, it is systematic for all data from a single pulse. Given this, and that it is typically at least as large as the difference between the three approximations, it is easy to find individual pulses which appear to agree with any of the models, as shown in figure 4.13.

As shown in section 4.2, due to the complex and partly unidentified nature of the variation in the calibration, the uncertainty on the measured ellipticity angle χ_d cannot at present be accurately characterised. However, over a suitably large range of pulses and campaigns, it can be seen that except for pulses with very large ellipticity, the predictions of the cold model agree on average with data-derived values from cold ($T_e <$ 3keV) pulses (see row 1 column 4 of figure 4.7). Under the assumption that the basic cold plasma propagation theory is correct for cold plasmas, this indicates that the average calibration is correct.

In this case, it is useful to investigate a large number of pulses over many years and a wide range of plasma parameters. Originally, all JET pulses between 2003 and 2007 with valid interferometry, polarimetry and LIDAR T_e data were selected, giving over 1200 pulses. For each of these, equation 3.42 was used to calculate an approximate prediction for χ which was compared against the data-derived χ_d for cold regions of each pulse (core $T_e < 3$ keV). Pulses with a large disagreement ($\chi_d - \chi > 20\%$) in these regions were rejected as having too large a calibration error to be useful, under the assumption that the the basic theory is correct for low T_e plasmas. In the high T_e regions of such pulses, the disagreement between the measurement and all the models would be much larger than the difference between the models. Also rejected were pulses with very high uncorrelated instrumental random noise and very large clear oscillations. This selection leaves 268 pulses which, at 250ms intervals, contain over 23000 data points.

4.3.4. Statistical Comparison

Figure 4.14 shows a statistical overview for the remaining pulses of the difference between data-derived ellipticity and that calculated using the plasma forward model with the three different forms of Ω.

The predictions from the cold model show a clear systematic underestimate for $\chi > 4^o$, the non-relativistic warm model shows a clear over-estimate and the relativistic model shows the least systematic disagreement. At first glance this seems fairly conclusive. However, even for high χ, any low T_e data points should be correctly predicted by the cold plasma approximation but almost no correctly predicted data can be seen for $\chi > 6^o$ in figure 4.14a. The absence of these could simply be because the high density and hence high χ JET pulses also tend to have high temperature but from this graph it cannot be ruled out that the average calibration simply gives an incorrect dependence on the real plasma χ, which is coincidentally well compensated for by the relativistic T_e correction.

4.3.5. Temperature Dependence

To demonstrate that the better agreement with the relativistic model seen in figure 4.14 is due to the high T_e effects, the disagreement with the cold plasma model must be shown to be correctly T_e dependent. A preliminary indication of some dependence can be seen in figure 4.14a at $\chi \approx 4^o$ where the high T_e points (red) do have, on average, a greater under-estimation

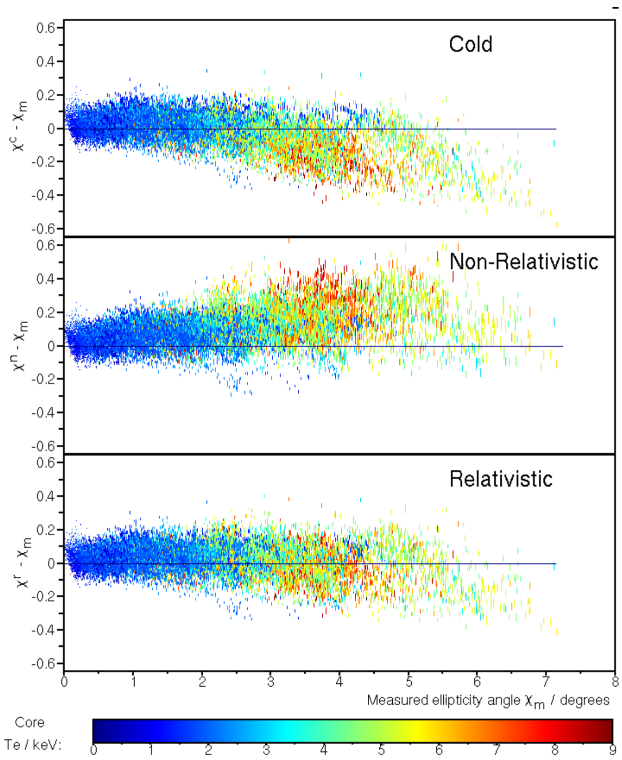

Figure 4.14.: Difference between calculated and data-derived induced ellipticity for 250ms intervals of 268 selected shots. Calculation performed with a) cold plasma model, b) non-relativistic and c) weakly relativistic linear T_e corrections. The 23000 data points are coloured by core T_e at each time point. Data point vertical height represents uncertainty in calculated χ from uncertainty in n_e profile and channel position.

than the medium T_e data (green). This cannot be seen in the relativistic model's predictions (figure 4.14c). To investigate this further, it is possible to examine the approximate dependence on T_e directly. The approximate ellipticity, in terms of the vector Ω, is given[53] by :

$$\chi \approx \tfrac{1}{2} \int \Omega_1 \, dz \tag{4.5}$$

For the cold plasma approximation, substituting Ω^c simply gives equation 3.42. Substituting Ω^n from equation 4.3 or Ω^r from 4.4 and assuming constant $B_\perp$ leads to the following expression for the difference between the respective warm approximation and the cold plasma model prediction:

$$\chi^{n,r} - \chi^c \approx \alpha^{n,r} \, C_1 B_\perp^2 \int \tau \, n_e \, dz \tag{4.6}$$

The constant α is $\alpha^n = 6$ for the non-relativistic model and $\alpha^r = 2.25$ for the relativistic model.

Figure 4.15 shows the difference between the measured and cold model ellipticity $\chi - \chi^c$ against $x = B_\perp^2 \int T_e \, n_e \, dz$ for all data points in figure 4.14 with $\chi \geq 1^o$. Those with $\chi < 1^o$ have been removed from the data set as they come almost exclusively from the pulse ramp up and down periods and do not contain any useful information about the temperature effects (see figure 4.3.4), leaving 15000 data points. Also shown are the linear relationships from equation 4.6 for both the non-relativistic and relativistic models and the $\chi - \chi_d = 0$ line representing the cold plasma approximation.

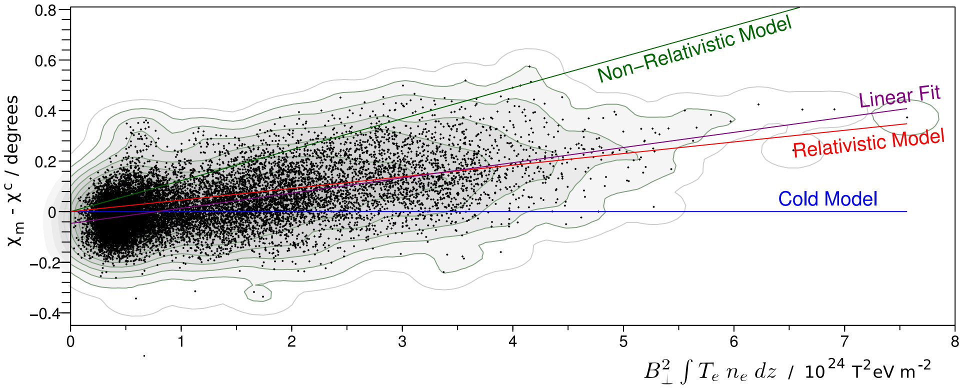

Figure 4.15.: Difference between measured ellipticity and that calculated from the cold plasma model $\chi - \chi^c$ plotted versus $x = B_\perp^2 \int T_e \, n_e \, dz$. Individual data points at 250ms intervals of 268 pulses are shown above contours of a kernel density estimate from these. Also shown are the approximate theoretical relationships for the three models and a simple linear least-squares fit to the data.

It is immediately clear that the data is in better agreement with the relativistic model and

	α	$\frac{\lvert\alpha-\alpha_{fit}\rvert}{\lvert\alpha^r-\alpha_{fit}\rvert}$	$\frac{\lvert\alpha-\alpha_{fit}\rvert}{\sigma_\alpha}$
Fit	2.94	-	-
Cold	0.00	4.26	59
Non-relativistic	6.00	4.43	62
Relativistic	2.25	1	14

Table 4.1.: Theoretical and fit values of α for equation 4.6; difference between each model and fit value normalised to the relativistic model and multiples of its uncertainty that fit value lies from each theoretical value for an estimate of $\sigma_\chi = 0.09^o$.

that qualitatively, the agreement with this model is good given the spread of the data by the calibration variation. This constitutes the experimental observation of relativistic effects from the JET polarimeter.

A simple linear least-squares fit to the data was performed assuming independent Gaussian distributed random noise of a constant value, denoted σ_χ, for all data. Without a complete model of the calibration uncertainty, σ_χ is difficult to assess but evaluated in terms of this, the linear fit gives a gradient of $\alpha_{fit} = 2.94$ with uncertainty $\sigma_\alpha = 0.57\sigma_\chi$ and with a constant offset of -0.05^o. The constant offset describes the part of the difference between the measurement and cold approximation which is not a function of temperature and most likely indicates a small bias is present in the average calibration.

Even without a specific value of σ_χ, it is possible to quantify the relative agreement of each of the three models by evaluating the relative difference of α_{fit} from α for each. Table 4.1 gives this difference for each model normalised to 1 for the relativistic model. Given that the non-relativistic and cold plasma models are over 4 times further from the fit value than the relativistic value, it can be seen quantitatively that the data is in far better agreement with the relativistic model.

Further to comparing the three models, it would be useful to assess to what degree the data agrees with the relativistic model. Unfortunately, α_{fit} is of little use for this without its uncertainty and therefore requires an absolute value for σ_χ. A very approximate estimate, found by analysing the spread of $(\chi - \chi^c)$ about the linear fit, gives $\sigma_\chi = 0.09^o$ and hence $\sigma_\alpha = 0.05$. Table 4.1 shows the multiples of this by which α_{fit} lies from the theoretical value, for each model. For the relativistic model, the fit gradient α_{fit} lies 14σ from α^r which appears to indicate that, while it out performs the others, the data does not actually support the relativistic model. This large disagreement comes from the apparently very small fit uncertainty

σ_α ($< 2\%$) which is not surprising given that it is based on 15000 data points. If the only uncertainty on the data was the independent Gaussian distributed random noise that was assumed, an average over this much data really would give a value for α_{fit} of this accuracy. Unfortunately, this is almost certainly not the case. For instance, since most of the variation is from the calibration uncertainty which is systematic for each pulse and the data only contains 268 pulses, the errors cannot be considered independent. To correctly assess the validity of the model, it is necessary to correctly and completely model the calibration procedure and the uncertainty in χ that results from it. With this, the quantity of data may even be sufficient to observe the finite T_e effects to second order in τ.

4.3.6. Re-calibration

While the comparison of the models in the previous section is almost entirely conclusive, it could be argued that the relatively subjective selection process of section 4.3.3 introduces a bias which favours the relativistic model. In reality, since the selection involved the comparison of the measurements to equation 3.42, it is most likely that it would bias the results toward the cold plasma model. However, further evidence can be seen for supporting the relativistic model from some of the the pulses rejected in section 4.3.4.

Pulses where the measured data completely disagrees with all the models can be re-calibrated by fitting an arbitrary $(R, R') \rightarrow (\psi, \chi)$ mapping so that the (R, R') signals map correctly to the predictions based on EFIT. For this analysis, again under the assumption that the theoretical model is correct for low T_e plasma, the signals are fit on periods of the pulse where $T_e < 3keV$ and are fit to the cold model. Figure 4.16 shows such a pulse with both its original calibration and after the calibration fitting procedure.

As the fit is performed to the predictions of the cold model, the procedure is most likely to bias the results toward this. Despite this, the re-calibrated measurement is in clearly better agreement with relativistic model. Figure 4.17 shows $\chi - \chi^c$ from the re-calibrated data along with the three theoretical approximations as in section 4.3.5. Also shown is the result of a linear least-squares fit to the data, which gives (as in section 4.3.5) a linear dependence of $\alpha_{fit} = 2.74 \pm 4.4\sigma_\chi$.

Again, it can be seen immediately that the agreement with the relativistic model is far better than with the other two. Quantitatively, the cold and the non-relativistic model values for α are respectively 5.6 and 6.7 times further from α_{fit} than the relativistic value. This gives further evidence to the main conclusion of section 4.3.5 and demonstrates that with accurate calibration, the JET polarimeter can observe relativistic effects on individual pulses.

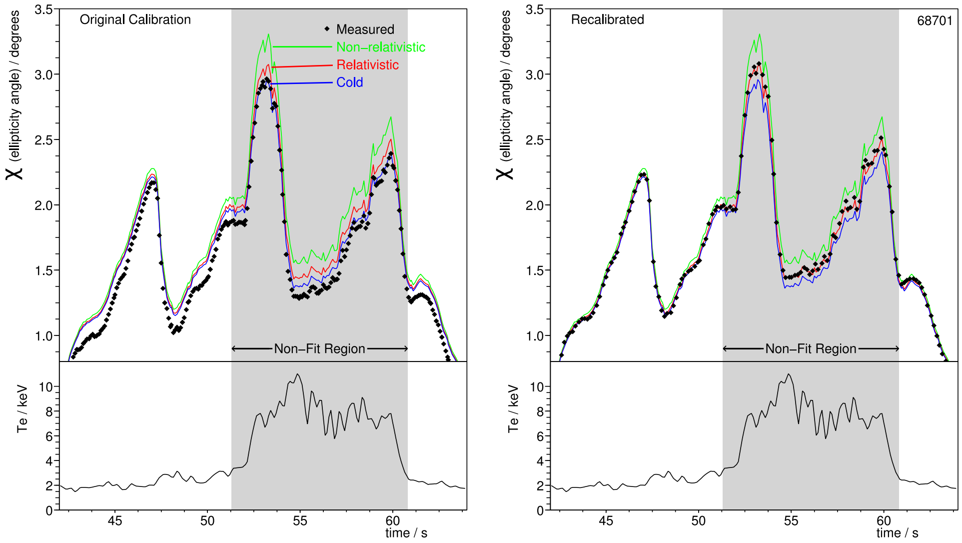

Figure 4.16.: Time trace for a single high T_e JET pulse showing calculated (lines) induced ellipticity χ from each model. Measured data (points) is shown using a) the original calibration and b) the calibration re-fit to the cold periods.

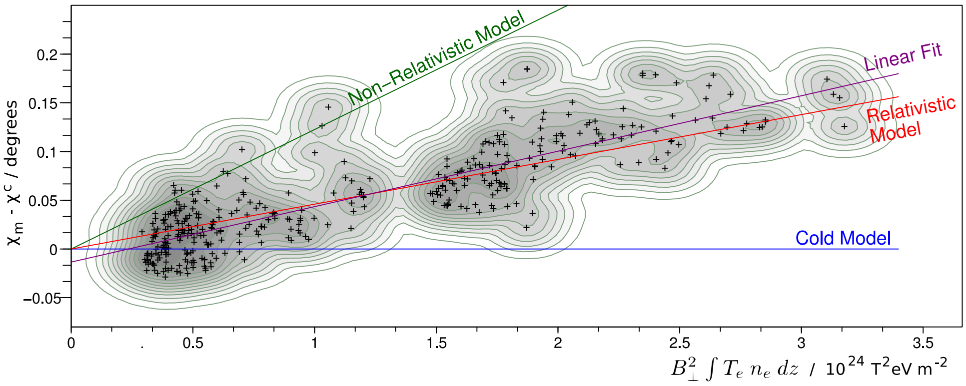

Figure 4.17.: Difference between measured ellipticity and that calculated from the cold plasma model $\chi - \chi^c$ plotted versus $B_\perp^2 \int T_e\, n_e\, dz$. Individual data points at 50ms intervals of a single pulse after re-calibration shown above contours of a kernel density estimate from these. Also shown are the approximate theoretical relationships for the three models and a simple linear least-squares fit to the data.

Unfortunately, the re-calibration procedure is restricted to a fairly small number of pulses as it requires a cold period $T_e < 3keV$ with χ of similar magnitude to the hottest period. As χ is approximately proportional to line integrated density, this requires part of the pulse to be of similar density but much lower temperature to the main heating phase. This is uncommon at JET where T_e and n_e typically both rise together before a single main heating phase after which they fall together.

4.3.7. Finite Temperature Effect Conclusions

A comparison has been made between experimental measurements of induced ellipticity from the core channel of the JET polarimeter and calculations based on the cold plasma approximation and including non-relativistic and also weakly relativistic finite T_e effects to first order in τ. To deal with difficulties with the instrument calibration, two methods were used. Firstly, a statistical view of pulses with a good original calibration was given and secondly, a single pulse was re-calibrated.

In both cases, it has been clearly shown that the measurements are in far better agreement with the relativistic model. Under the assumption that the basic theory is correct (i.e the cold plasma model is valid for cold plasmas), this verifies the presence of relativistic effects in the propagation of high frequency radiation through high T_e plasma. It demonstrates that both non-relativistic and weakly relativistic finite T_e effects must be taken into account when considering ellipticity measurements of plasmas with $T_e \sim> 5keV$.

While a more detailed model of the calibration would be required for a direct and complete verification of the model's validity, both methods give an approximate T_e dependence relatively close to the relativistic finite T_e model's prediction.

5. LIDAR Thomson Scattering

Along with the magnetic field and electron density n_e, one of the most important plasma parameters is the electron temperature, T_e. At the time this project begun, the two principal diagnostics which gave information on T_e were the Electron Cyclotron Emission and the Core and Edge LIDAR Thomson Scattering systems. The two LIDAR systems have sufficiently similar design that a single model could be developed for both systems and given that they each also provide information on n_e, they could be coupled with the Interferometry model from chapter 3. For these reasons, it was decided to include the two LIDAR systems and this chapter details the modelling and analysis of both.

The model of the physics involved, covered by section 5.1 has been developed previously and is relatively simple to implement but the diagnostic system itself, is also highly complex. For the inference of T_e and n_e to be accurate, and to extract the maximum possible information from the data, it was necessary to model this complexity in extreme detail and to include all the calibration parameters in the forward model. Sections 5.2 through 5.4 detail the full operation and existing calibration procedures of both diagnostics as determined from the original documentation and discussions with the diagnostic operators. The remaining sections account work carried out entirely by the author. Section 5.5 describes the development of the full likelihood function and some approximations of it. Section 5.6 describes the work done to determine and improve the accuracy of the calibration information and the effect it has on the inference of n_e and T_e. Section 5.7 gives some typical results for JET H-mode plasmas and shows the improvement gained over the standard analyses by using the full models. Finally, section 5.8 attempts to use the new models to examine the shape and evolution of the H-mode T_e and n_e pedestals.

5.1. Thomson Scattering

The process of *Thomson scattering (TS)* is exploited in many plasma experiments to obtain measurements of both ion and electron temperatures and densities. The measurements are made by analysing the spectrum and intensity of light scattered from a high power monochromatic laser pulse sent through the plasma. On JET, three systems use the Thomson Scattering principle: The Core[66] and Edge[67] LIDAR systems and the more recent High Resolution Thomson Scattering[26] system. All three systems are employed for measuring n_e and T_e profiles across different regions of the plasma.

The scattering of light by a charged particle can be viewed within quantum mechanics as the collision of the particle and a photon. If the photon energy $\hbar\omega$ is comparable to the rest mass of the particle, the collision can change its momentum and the process is known as Compton scattering. If the charged particle momentum is almost unaffected, the process is known as Thomson scattering and can be viewed classically as the electron being accelerated by the incident wave and emitting radiation as result. The lasers in both JET LIDAR systems have a wavelength $\lambda_l = 694nm$, leading to photons with an energy of $2eV$ which, at six orders of magnitude smaller than the rest mass of an electron, makes Thomson Scattering a valid approximation.

The incident radiation will accelerate both ions and electrons but the small electron to ion mass ratio ensures that the induced ion velocities are negligible and so only electron scattering is usually detected. The Thomson scattering process is usually split into two cases depending on how the wavelength of the incident light compares to the plasma's Debye length λ_D. If the incident wavelength is much longer $\lambda_l >> \lambda_D$, the shielding charge surrounding each electron oscillates with opposite phase to that electron, cancelling out the radiation. While ions do not themselves radiate significantly, scattering from their shielding electrons is detected in this case, meaning the scattered light holds information about the ions. This is known as 'coherent scattering'. If the incident wavelength is much shorter than the Debye length $\lambda_l << \lambda_D$, the scattering from an electron's shielding cloud does not cancel that electron's radiation and so information on the electron is detected, this is known as 'incoherent scattering'. For JET plasmas, $\lambda_D \approx 10^{-4}m$, far longer than the laser wavelength and so incoherent scattering is observed in both LIDAR systems.

The spectrum of incoherent scattering arises because of the Doppler effect. Each electron re-radiates light in its rest frame at the incident frequency it sees in its rest frame. For an electron travelling toward the laser, the electron sees a shorter wavelength (blue-shifted). The emitted radiation, as seen by an observer in the laser's direction (and rest frame) is seen as further blue-shifted due to electron's velocity. If seen by an observer in the opposite direction it will be red-shifted back toward the incident frequency. It is clear then, that the spectrum observed will depend on the electron velocity distribution. For the vast majority of scenarios, the tokamak plasma is in thermodynamic equilibrium making the distribution Maxwellian. For temperatures above $1keV$, which JET plasma usually are, the high velocity tail contains electrons travelling at relativistic speeds and a relativistic treatment of the Doppler shift must be used. The full derivation of this can be found elsewhere [68]. For temperatures of order 30keV, which are beyond JET's operating range but will be present in ITER, the next large tokamak to be built, the relativistic effects also include a significant change to the polarisation and the scattering becomes difficult to calculate exactly. Fortunately, a sufficiently accurate approximation is given elsewhere[69], suitable for 644nm TS systems and any plasma across the full operating range of both JET and ITER. The formula is shown below for a fixed scattering angle of 180^o as both LIDAR systems detect almost entirely backscattered light. It gives the approximated scattered power P_s per unit solid angle Ω_s per unit normalised wavelength ϵ, where λ_s is the scattered wavelength, λ_l is the laser wavelength, P_i is the total power incident on the scattering volume containing n_e electrons and r_e is the classical electron radius:

$$\begin{aligned}
\frac{d^2P_s}{d\Omega_s d\epsilon} &= r_e^2 n_e P_i \frac{e^{-2\alpha}q}{2K_2\left(2\alpha\right)\left(1+\epsilon\right)^3\sqrt{4\left(1+\epsilon\right)+\epsilon^2}} \qquad (5.1)\\
q &= 1-4\eta\frac{2+\eta}{2+13\eta} \qquad \eta=\frac{1}{2\alpha x}\\
x^2 &= 1+\frac{\epsilon^2}{4\left(1+\epsilon\right)} \qquad \epsilon=\frac{\lambda_s-\lambda_l}{\lambda_l} \qquad 2\alpha=\frac{mc^2}{T_e}
\end{aligned}$$

Figure 5.1 shows the Thomson scattering spectrum as calculated by equation 5.1 for a range of temperatures typical of JET plasmas along with the spectral sensitivity profiles of each channel of the Core and Edge LIDAR spectrometers. To determine the detected signal in each channel, the number of photons scattered per second that are in the spectral range of that channel must be calculated. With $R(\lambda_s)$ as the spectral response function the channel, this is:

$$\frac{dN_s}{dt} = \int \frac{d^2N}{dt\, d\lambda_s} R(\lambda_s) d\lambda_s \qquad (5.2)$$

The power per unit wavelength in equation 5.1 can be converted via the photon energy

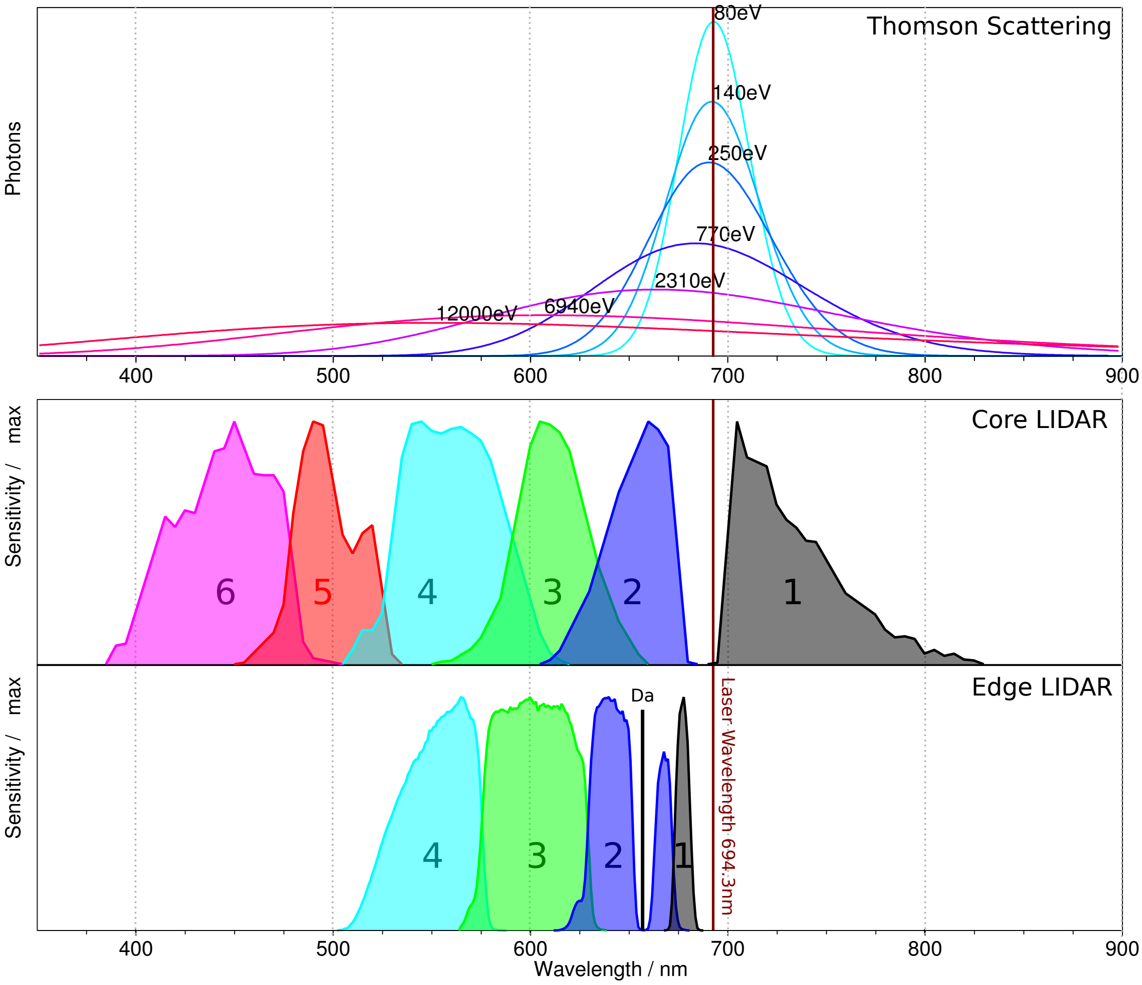

Figure 5.1.: a) Thomson backscattering spectra $dP/d\lambda$ from $694.3nm$ incident light at a range of temperatures. b) Normalised detector spectral response profile $R(\lambda_s)$ for the 6 Core LIDAR spectral channels and c) for the 4 Edge LIDAR spectral channels.

expression $E = hc/\lambda$ to the total number of photons per unit wavelength per unit time:

$$\frac{dN_s}{dt} = \int \frac{\lambda_s}{hc} \frac{d^2P}{dt\, d\lambda_s} R(\lambda_s) d\lambda_s \tag{5.3}$$

The predicted count from one channel in a time period Δt for a given n_e and T_e, assuming the photons are emitted as a constant stream, is:

$$N_s\left(n_e, T_e\right) = \Delta t \int \frac{\lambda_s}{hc} \frac{d^2P}{dt\, d\lambda_s} R(\lambda_s) d\lambda_s \tag{5.4}$$

Because the photons are emitted randomly, the count that would actually be observed must be expressed as a probability distribution which is the likelihood function for this simple case. This is the Poisson distribution, where M is the number of photons actually observed:

$$P\left(M|N_s\right) = \frac{N_s^M e^{-N_s}}{M!} \tag{5.5}$$

When the photon counts are significantly above around 30, this can be approximated as a Gaussian distribution with $\sigma = \sqrt{N_s}$:

$$P\left(M|N_s\right) \approx \frac{1}{\sqrt{2\pi N_s}} exp - \frac{\left(M - N_s\right)^2}{2N_s} = \mathcal{G}\Big(M; N_s, \sqrt{N_s}\Big) \tag{5.6}$$

The usual analysis procedure, which is adopted by the JET standard analysis, is to fit the prediction N_s to the detected photon count using this likelihood function. An important assumption in this fitting procedure is that all of the light which contributes to a single reading comes from plasma of the same n_e and T_e. Significant variation within the collection region of either, e.g. sharp gradients, will result in a spectrum which is the sum of different spectra. Temperature gradients particularly, can result in detected spectra which are not even consistent with equation 5.1 for any single T_e. In most Thomson scattering systems this assumption is valid since the collection optics are each focused on a single small volume and each point is measured by a different set of optics. The LIDAR systems, however, use a very different approach.

5.2. LIDAR

The JET Core and Edge LIDAR systems are based on the *Light Detection and Ranging (LIDAR)* principle. In each measurement *frame*, a 300ps laser pulse is sent through the plasma and a single set of collection optics and detectors measure the light back-scattered along the laser line. If the detectors and laser were in the same place, the position along the line of sight from which a photon has scattered is determined by its time of flight $z = c\left(t_D - t_L\right)/2$ where t_D is the time of arrival at the detectors and t_L the time the laser pulse left the laser. This simple situation is shown in figure 5.2a. The fitting procedure described previously can be carried out for the light detected at a series of times, each giving the T_e and n_e at the corresponding z. The complete result would be a profile of both variables along z. Because each volume of plasma is only illuminated for the $\sim 300ps$ it takes the light pulse to pass through it, the resulting profiles are an incredibly short 'snapshot' of the plasma, compared to conventional TS systems which typically integrate each region over at least a few $10ns$. The whole procedure is then performed throughout the plasma pulse, 140 times at 4Hz for the Core LIDAR system at 6 times at 1Hz for Edge LIDAR system.

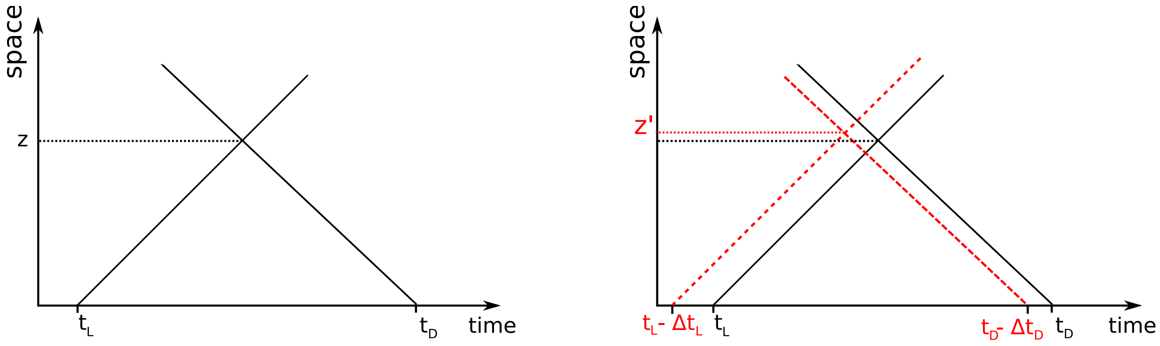

Figure 5.2.: Space-time diagrams for the LIDAR system showing a) a single photon leaving the laser at t_L, scattering at z and returning at t_D and b) with an extra scattering from a different position z' that could be included with the first. The vertical axes represents the distance along the beam-path on which the laser and detectors both are positioned at $z = 0$.

In practice, the situation is complicated firstly by the finite width of the laser pulse. It is possible that light arriving at the detectors at t_D was actually scattered from light that left the laser a short time Δt_L before or after the central laser firing time t_L. The relative intensity of light emitted by the laser at different Δt_L is given by its pulse shape $L\left(\Delta t_L\right)$. The second complication is the finite time resolution of the detection system. It is possible that the signal measured at t_D actually came from light which arrived a short time Δt_D before this. The relative sensitivity of the complete detection system is given by $D\left(\Delta t_D\right)$. Figure

5.2b shows the general case for an arbitrary Δt_D and Δt_L where the scattering takes place at $z' = c\left(t_D - t_L - \Delta t_D + \Delta t_L\right)/2$. The total number of photons collected at any t_D is then the integral of the scattered photons $N_s\left(z'\right)$ from all such contributions, weighted by $L\left(\Delta t_L\right)$ and $D\left(\Delta t_D\right)$ and is given in equation 5.7.

$$N_{\gamma 1}\left(t_D|t_L\right) = \int_{-\infty}^{\infty} L\left(\Delta t_L\right) \int_{-\infty}^{\infty} D\left(\Delta t_D\right) N_s\left(c\frac{t_D - t_L + \Delta t_L - \Delta t_D}{2}\right) d\Delta t_D \, d\Delta t_L \qquad (5.7)$$

This double integral is a convolution of $N_s\left(z\left(t\right)\right)$ with both L and D and can be rewritten as $N_s \otimes C$ where C is the combined system resolution function $C = L \otimes D$. The width of the function C determines the length along the line of sight of the volume from which the spectrum detected at any instant is collected. It is this width which must be examined, for a LIDAR TS system, to validate the assumption of constant n_e and T_e over the scattering volume. The naive point-by-point Thomson spectrum fitting carried out by the standard analysis will have an effective spatial resolution of this length.

For the JET Core LIDAR diagnostic, L is a Gaussian with $FWHM \approx 300ps$ and D is similar to a Gaussian with $FWHM \approx 700ps$. This gives a convolution function with $FWHM \approx 760ps$ which means scattering is collected from a region of approximately 12cm. For the core of a stable JET plasma with no transport barriers, MHD activity or similar, n_e and T_e will not vary significantly over this scale. This is the regime for which the system was originally designed and for which the standard fitting procedure is sufficient. However, when transport barriers are present, especially at the edge, both n_e and T_e will simultaneously change rapidly over as little as 2cm. It is clear that in this case the assumption is broken and the simple fitting procedure does result in serious systematic errors for both profiles.

For the Edge LIDAR system, L and D have $FWHM \approx 300ps$ and $550ps$ giving C with $FWHM \approx 630ps$ and a scattering volume of 9cm along its line of sight. Despite the fairly shallow angle between the line of sight and the flux surfaces, the H-mode pedestal projected onto the line of sight is still usually < 10cm and so proper consideration of the effective convolution is essential when using Edge LIDAR data from H-Mode plasmas.

5.3. Optics and Timing

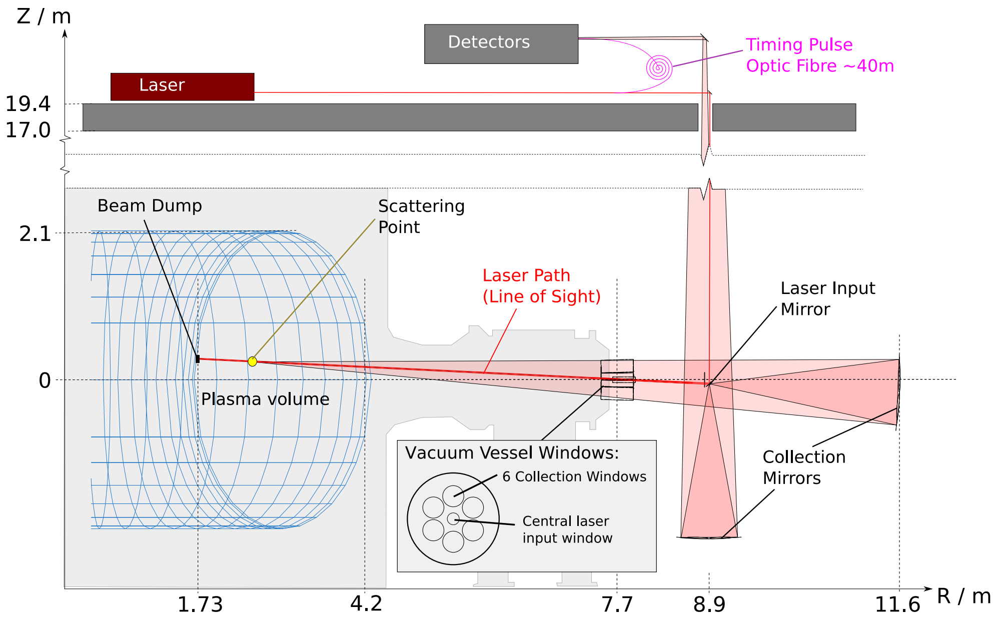

Figure 5.3.: The Core LIDAR Thomson Scattering system optical setup showing scattering from an arbitrary position along the system's line of sight. The plasma, windows and collection mirrors are to scale.

The layout of the input collection optics for the JET LIDAR systems are shown in figures 5.3 and 5.4. In both cases, the laser and spectrometer are situated in the roof laboratory 20m above the vessel. For the Core LIDAR system, a series of mirrors direct the laser through the central vessel window and across the plasma. The backscattered light passes through 6 window tubes located around the input window and is delivered to the spectrometer in the roof lab by means of the 2 large sets of collection mirrors and a small mirror near the laser input mirror, all of which are not mounted on the Tokamak itself. The Edge LIDAR system is similar, but has collection optics physically mounted on the vessel. The solid angle of collection is less for smaller major radius R which leads to a reduced sensitivity toward the very inboard side of the plasma. The very outboard edge also has reduced sensitivity due to details of the setup of the collection optics and spectrometer. The amount of light collected by the detectors as a function of R is known as the system's vignetting curve. It is measured by filling the vessel with a constant density of Nitrogen, firing the laser and measuring the Raman scattering[70]. This scattering should be equal everywhere along the line of sight and so the measured intensity variation should be this vignetting curve. The result of this calibration is shown in figure 5.5 for both systems.

The input and return path lengths are very different which means the simple relationship

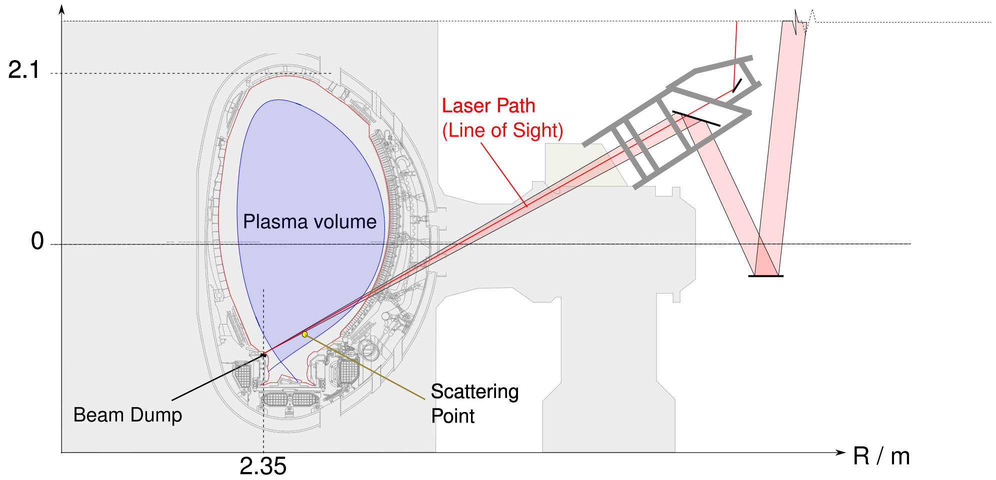

Figure 5.4.: The Edge LIDAR Thomson Scattering system optical setup showing scattering from an arbitrary position along the system's line of sight.

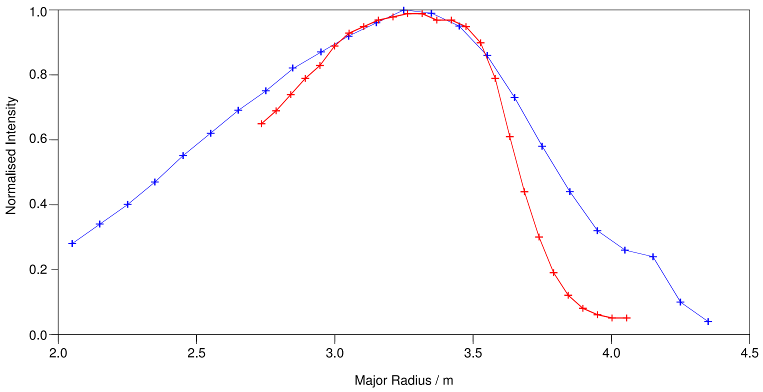

Figure 5.5.: The measured vignetting calibration curve (2006) for The Core LIDAR system (blue) and the Edge LIDAR system (red). The curve describes the relative amount of light collected from different radial position.

$z = c(t_D - t_L)/2$ cannot be used to determine the scattering positions. Instead, each system is calibrated to its laser beam-dump, the position of which is better known. As the laser pulse hits the beam dump a large reflection pulse is emitted back toward the collection optics. The time this light arrives at the detector provides a reference point from which to draw the position of the other detected scattering. Unfortunately, this 'back-wall pulse' is orders of magnitude greater than the Thomson scattered light for which the detectors are optimised and so they are completely saturated by it. To resolve this problem, an approximately 40m loop of optic fibre redirects a small portion of the laser light directly into the detectors. The length of the fibre is sufficient for this 'timing pulse' to arrive at the detectors shortly before light from the plasma scattering arrives. In a separate calibration experiment, in which the detector gain is turned down sufficiently to see the back-wall pulse clearly, the time difference between receiving the timing and back-wall pulses is measured. In normal scattering measurements, the timing pulse is used as the time point from which the position of the TS light is deduced. The procedure is actually only valid for the first channel of each spectrometer since only that will see reflected light. It is believed the other channels see light scattered from a small ablation plasma created as the pulse hits the beam dump, which will be slightly delayed. To work around this problem, the difference in path length between the other detectors and the first is also measured directly with a ruler. For the Edge LIDAR system, the detectors are positioned so that they all have exactly the same path length from the beam dump.

While all of these measurements theoretically allow the position to be deduced, the calibration procedure is sufficiently complex for both systems that significant uncertainty in the position of the profiles may be introduced. The calibration measurements are made at most once every few months and any variation in these periods of the optics or detectors will not be known (hence not modelled) and will introduce systematic errors that will not be accounted for. It will later be shown that correctly modelling the LIDAR positioning and timing is vital to their correct analysis and that with the Bayesian integrated approach, none of these calibrations are entirely necessary.

5.4. Detectors

The LIDAR detector systems consists of a series of filters which reflect successively shorter wavelength portions of the scattered spectrum into Micro-channel plate *photomultiplier tubes (PMTs).* Each PMT has a photocathode, which converts a fraction of the incident photons into electrons (known as photoelectrons). This fraction is known as its quantum efficiency (QE). Each photoelectron is accelerated through a series of plates of increasingly high voltage. The impact of any electron with each plate releases more electrons resulting in a large number eventually reaching the PMT's anode, from which a current pulse is generated. This pulse is fed to an *Analog to Digital Converter (ADC)*, where, after some electronic signal processing the voltage level is converted to a digital number and stored. This digital value is the data which is predicted by the forward function developed here. In order to do this, the model must include the details of the optics, PMT and ADC behaviour for all channels of both systems.

In both the optics diagrams (figures 5.3 and 5.4), it can be seen that the laser pulse passes through a vessel input window shortly before entering the plasma. Since the input windows are not perfectly transparent, a small fraction of the pulse is reflected. Similarly to the beam-dump reflection, this is much greater than the TS scattering and the large number of photo-electrons created in the PMTs would saturate the detection until long after the timing pulse and TS light from the plasma has arrived, meaning nothing could be measured. To avoid this, the high-voltage supplies to the PMTs are not switched on until shortly before the timing pulse arrives.

Unfortunately, switching on the HV supplies causes significant electronic oscillations to appear. These oscillations, while complex, appear in the data with the same shape and amplitude for every frame on each channel regardless of the light arriving at the PMT and hence are known as the *baseline.* Just before each JET plasma pulse, the detector system switches on the PMT HV supplies and records one complete frame of data without the laser firing. These frames show only the baseline signal and so that frame can be included in the forward function to simulate the baseline on the real plasma frames.

When the laser is fired and/or there is a plasma in the vessel there is a second effect of the HV switch-on. While the HV supply is off, both background plasma light and reflection from mirrors and the input window incident on the photo-cathode still generate photoelectrons. As these are not accelerated away they collect at the photo-cathode and when the switch-on occurs are accelerated in one go, giving a single large pulse known as the *switch-on pulse.* Unlike the baseline, the amplitude of this pulse is heavily dependent on the specific conditions at the time of the frame. After the switch-on pulse has subsided, the signal returns to the baseline level but with an extra constant positive offset resulting from any background/ *ambient light* flux received from the plasma. This level is constant because the significant background emission from the plasma does not fluctuate significantly over the frame capture period.

Figure 5.6 shows the typical signal for a complete frame of the Core LIDAR system which results from all these components. The Edge LIDAR system is very similar in construction but has a significantly larger baseline and often negligible switch-on pulse. Figure 5.7 shows a schematic diagram of the whole system together with the flow of the information through the system and their notation which will be used in this document.

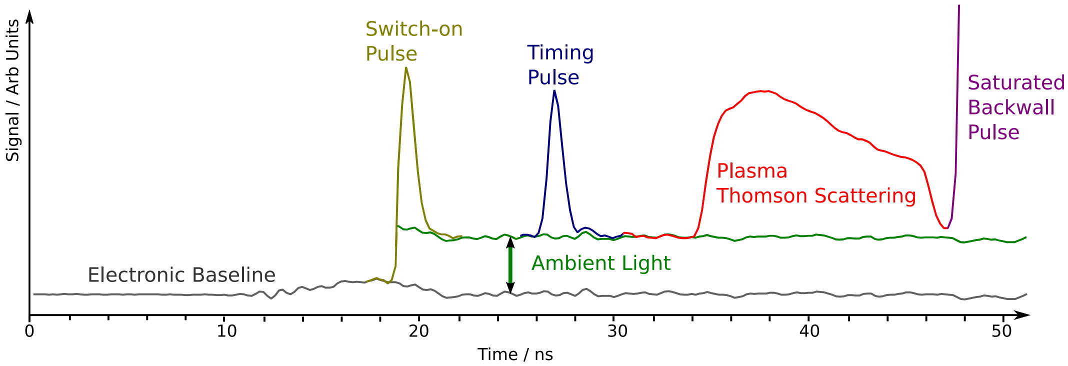

Figure 5.6.: A typical Core LIDAR signal from one detector for a single frame.

It can be seen from figure 5.7 that the overall signal intensity, while proportional to the electron density, depends heavily on many factors such as transmission of the windows, the optical path and all of the PMT characteristics. Measuring these individually and even measuring the complete absolute transmission is difficult. Instead, the *relative sensitivity* of each channel Λ_{ch} is determined by a calibration experiment and the *overall transmission coefficient* Λ_* is left as a free parameter. This means the absolute electron density cannot be determined by these diagnostics. A third set of parameters ξ_{ch} are given to the proportionality of the final data D to primary photo-electron count $N_{\gamma e^-}$ per ADC sampling period, on each channel. These values depend on the high-voltage applied to the PMT anodes and are determined by the diagnostic team during a separate calibration experiment.

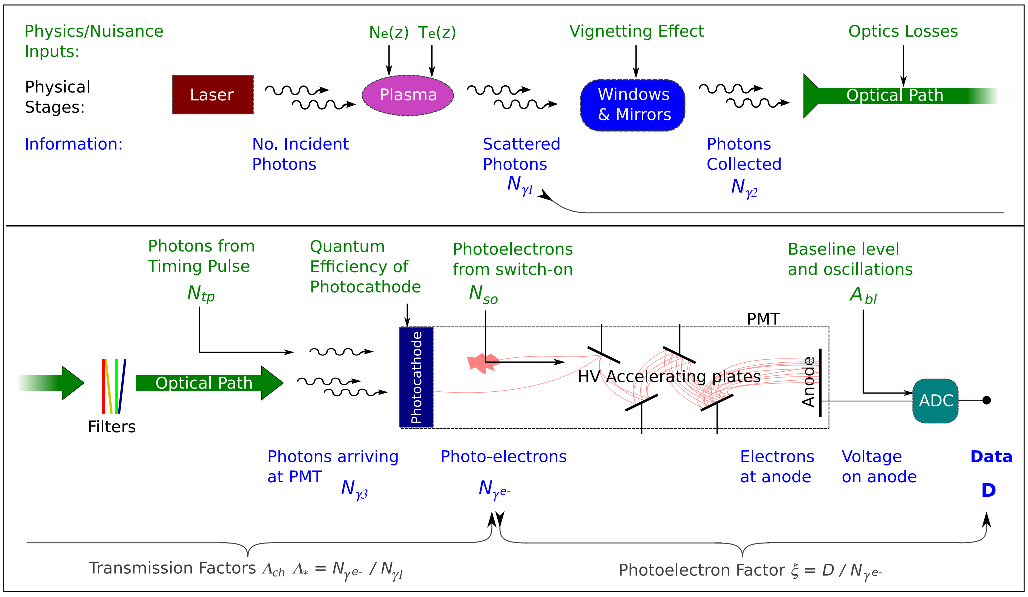

Figure 5.7.: A schematic representation of the equipment/processes of the LIDAR system. Below them (blue) are shown the information flow through the system and above (green) are shown the inputs and nuisance parameters that are used in the forward function.

5.5. The LIDAR Likelihood Distribution

5.5.1. Basic Foward Function

Construction of the complete forward model for the LIDAR system, to predict the data given a complete set of physical and calibration parameters, follows figure 5.7. The exact procedure is straight forward, though not trivial:

- Find n_e and T_e at series of distances $\underline{\text{x}}$, along the line of sight.
- Calculate arrival time of each point's information at the ADC $\underline{\text{t}}$
- Calculate scattered photon count per incident photon count into each channel's spectral range, at each point $N_{\gamma 1}(t, ch)$.
- Multiply by relative and overall sensitivity $\Lambda_{ch}\Lambda_*$ to get photo electrons arriving at each channel's spectrometer from TS light.
- Add ambient light, switch-on pulse, timing pulse and stray-light contributions to get total photo-electron count $N_{\gamma_{e-}}(t, ch)$.
- Convolve with combined instrument function C and multiply by photo-electron factor to get data contribution from light $D_{li}(t, ch)$.
- Add data contributions from electronic background to give final predicted data D_{pred}.

For the LIDAR system, the largest source of uncertainty is that relating to whether or not a photo-electron will be created when a photon hits the photo-cathode. Even given a constant stream of photons arriving at regular intervals, the PDF of the number of photo-electrons emitted in any period $P(N^0_{\gamma_{e-}} \,|\, \underline{\mu})$ will follow the Poisson distribution about the mean number over many such periods $N^0_{\gamma_{e-}}$. Conceptually similar uncertainties also arise from other sources like the probability of the Thomson scattering itself, or of electrons causing cascades in the PMT but all of these involve much larger numbers than the primary photo electron count. As this is the place in the system when the important information is held in the smallest discrete number, it contributes the largest uncertainty. To calculate the likelihood function $P(D_{pred} \,|\, \underline{\mu})$ requires performing the last four stages of the above to $P(N_{\gamma_{e-}} \,|\, N^0_{\gamma_{e-}})$ rather than just to $N^0_{\gamma_{e-}}$

5.5.2. Effect of the convolution on the likelihood.

Even assuming the Gaussian approximation (equation 5.6), the propagation of this PDF through the convolution with C [1] is especially complicated and three approximations were developed as part of this work. In the discussion of these that follows, the mean number of photoelectrons $N^0_{\gamma e-}$ will be referred to simply as $\underline{\mathrm{N}}$, and the number actually observed $N_{\gamma_{e-}}$ as $\underline{\mathrm{M}}$. The vector nature represents the fact that each count is over one ADC sampling period and the resulting vector relates directly to the vector of data observed over one frame.

The 'Simple' Approximation

If both the laser and detector were ideal (delta functions) and the ADC response a top-hat function of exactly 1 period width, there would be no convolution and the light contribution to each data point would come from counting photo-electrons only during that exact ADC period. Each data point would be entirely independent and the full likelihood distribution simply the product of the likelihood for each data point (equation 5.8).

$$P\left(\underline{\mathrm{M}} \mid \underline{\mathrm{N}}\right) \propto \prod_{i=1}^{n} \mathcal{G}\Big(M_i;\;\; N_i, \sqrt{N_i}\Big) \tag{5.8}$$

It is a likelihood function of this form that is effectively assumed in the standard analysis of LIDAR at JET and the approximation causes the inference of features much smoother and broader than in reality.

The 'Convolved' Approximation

To include the convolution, it is tempting to convolve the expected counts $\underline{\tilde{\mathrm{N}}} = \underline{\mathrm{N}} \otimes \underline{\mathrm{C}}$ and apply the Poisson likelihood to the result, ignoring the effect of the convolution on the distribution itself:

$$P\left(\underline{\mathrm{M}} \mid \underline{\mathrm{N}}\right) \propto \prod_{i=1}^{n} \mathcal{G}\left(M_i;\;\; \tilde{N}_i, \sqrt{\tilde{N}_i}\right) \tag{5.9}$$

This leads to a dramatic over estimation of the uncertainty. To understand why, consider a vector of constant A that is convolved with a normalised convolution function $\underline{\mathrm{C}}$. The result is the same constant level:

$$\underline{\mathrm{B}} = \underline{\mathrm{A}} \otimes \underline{\mathrm{C}} = A \tag{5.10}$$

---

[1] Technically, the convolution with the laser pulse function $L(t)$ should be applied before the appearance of the PDF as only the detector and ADC convolutions really happen after the source of uncertainty. Since only the complete function $C = L \otimes D$ has been determined experimentally and D makes up the larger part of it for both systems, this is ignored in this work. This will not effect the likelihood mean at all and will cause only a very small underestimation of its width and overestimation of the expected correlation.

However, if some Gaussian noise of constant level σ_A is added to the original signal, the noise level on $\underline{\mathrm{B}}$ is not σ_A. Rewriting the convolution as a sum, the effect on the standard deviation of the Gaussian noise can be found easily by Gaussian error propagation:

$$\begin{aligned} B_i &= \sum_{j=0}^{n} C_j A_{i-j} &\quad (5.11)\\ \sigma_B^2 &= \sum_{j=0}^{n} C_j^2 \sigma_A^2 = \sigma_A^2 \sum_{j=0}^{n} C_j^2 &\quad (5.12)\end{aligned}$$

The reduction in variance η, for an arbitrary instrument function $\underline{\mathrm{C}}$ or for a Gaussian of width σ_C is:

$$\eta^2 = \frac{\sigma_B^2}{\sigma_A^2} = \sum_{j=0}^{n} C_j^2 = \frac{1}{2\sigma_C\sqrt{\pi}} \tag{5.13}$$

The result $\underline{\mathrm{B}}$, is simply a smoothed copy of $\underline{\mathrm{A}}$ and the longer the convolving function, the smoother the signal. Applying this correction to the likelihood function of equation 5.9 results in equation 5.14.

$$P\left(\underline{\mathrm{M}}|\underline{\mathrm{N}}\right) \propto \prod_{i=1}^{n} \mathcal{G}\Big(M_i;\ \tilde{N}_i,\ \eta\sqrt{\tilde{N}_i}\Big) \tag{5.14}$$

This approach is used as a simplification later for ease of computation. While it takes into account the effect of the convolution on the magnitude of each point, it does not account for the large correlations introduced into neighbouring data points, which is not reflected in this PDF. An alternative way of viewing the flaw in this reasoning is that it is really $\underline{\mathrm{M}}$, the realised counts, that is convolved and not $\underline{\mathrm{N}}$, the expected counts.

The full 'Correlated' calculation

The effect of the correlations can be included by describing the distributions with a single multivariate Gaussian with each dimension representing the count within one ADC period. In this formalism, the approximated Poisson distribution (which is uncorrelated) for all counts is written as:

$$\begin{aligned} P\left(M_i \mid N_i\right) &= \mathcal{G}\Big(M_i;\ N_i, \sqrt{N_i}\Big) &\quad (5.15)\\ P\left(\underline{\mathrm{M}} \mid \underline{\mathrm{N}}\right) &= \mathcal{G}\big(\underline{\mathrm{M}};\ \underline{\mathrm{N}}, \underline{\underline{\sigma}}\big) &\quad (5.16)\end{aligned}$$

With $\sigma_{ij} = \sqrt{N_i}$ for $i = j$ and 0 for $i \neq j$, or $\underline{\underline{\sigma}} = \mathrm{diag}(\underline{\mathrm{N}})$. The convolution is really performed on the vector of the counts actually obtained $\underline{\mathrm{M}}$ and can be written as a matrix operation:

$$
\begin{aligned}
\underline{\tilde{\mathrm{M}}} &= \underline{\mathrm{M}} \otimes \underline{\mathrm{C}} && (5.17)\\
\tilde{M}_i &= \sum_{j=0}^{m} C_j M_{(i-j)} = \sum_{k=i-m}^{i} C_{(i-k)} M_k && (5.18)\\
\underline{\tilde{\mathrm{M}}} &= \underline{\underline{\mathrm{C}}}\,\underline{\mathrm{M}} && (5.19)
\end{aligned}
$$

The convolution matrix is then $C_{ik} = C_{i-k}$. The desired likelihood distribution is $P\left(\underline{\tilde{\mathrm{M}}} \mid \underline{\mathrm{N}}\right)$ which can be obtained by transformation of the Gaussian in equation 5.16 by $\underline{\underline{\mathrm{C}}}$. The new Gaussian has mean $\underline{\underline{\mathrm{C}}}\,\underline{\mathrm{N}}$ and covariance $\underline{\underline{\mathrm{C}}}\,\underline{\underline{\sigma}}\,\underline{\underline{\mathrm{C}}}^T$:

$$
P\left(\underline{\tilde{\mathrm{M}}} \mid \underline{\mathrm{N}}\right) = \mathcal{G}\left(\underline{\tilde{\mathrm{M}}};\ \underline{\underline{\mathrm{C}}}\,\underline{\mathrm{N}},\ \underline{\underline{\mathrm{C}}}\,[\mathrm{diag}(N)]\,\underline{\underline{\mathrm{C}}}^T\right) \qquad (5.20)
$$

This gives the probability of obtaining any vector of counts $\underline{\tilde{\mathrm{M}}}$ after the detector's convolving effect, given the expected vector of counts $\underline{\mathrm{N}}$. The principal difference between this and equation 5.14 is that this PDF assigns a low likelihood to data that has changes faster than the convolution width and larger than the electronic noise. Drawing samples from equation 5.14 will produce data with excessive high frequency noise. For inversion of real data, neither the data or the mean prediction can ever have such changes since both come from the convolution of a noisy signal, with only electronic noise added afterwards. Almost no set of parameters can produce a prediction ($\underline{\underline{\mathrm{C}}}\,\underline{\mathrm{N}}$) which involves evaluating the likelihood where the approximation and full calculation differ significantly. Some inversions were performed using both models and gave little difference in the resulting posterior PDFs. Given this, and the fact that equation 5.14 is significantly less computationally expensive, the convolved approximation is used for the results in the following sections.

5.5.3. Electronic Baseline and Noise

The only uncertainty which is comparable to that due to the uncertain number of photo-electrons is the electronic noise picked up between the PMT and ADC - σ_{elec} in Data/ADC units. Since the ADC convolution is not much more than 1 ADC point, this can be treated a simple uncorrelated Gaussian likelihood and be added to each PDF. Included this, converting from photo-electrons to full ADC units via ξ and adding the electronic baseline level A_{bl}, the three models are:

$$
\begin{aligned}
P\left(\underline{\mathrm{D}}\,|\,\underline{\mathrm{N}}\right)^{\mathrm{simp}} &\propto \prod_{i=1}^{n} \mathcal{G}\Big(D_i;\ \xi N_i + A_{bl},\ \sqrt{\xi^2 N_i + \sigma_{elec}^2}\Big) && (5.21)\\
P\left(\underline{\mathrm{D}}\,|\,\underline{\mathrm{N}}\right)^{\mathrm{conv}} &\propto \prod_{i=1}^{n} \mathcal{G}\Big(D_i;\ \xi \tilde{N}_i + A_{bl},\ \sqrt{\xi^2 \eta^2 \tilde{N}_i + \sigma_{elec}^2}\Big) && (5.22)\\
P\left(\underline{\mathrm{D}}\,|\,\underline{\mathrm{N}}\right)^{\mathrm{corr}} &\propto \mathcal{G}\Big(\underline{\mathrm{D}};\ \xi \underline{\tilde{\mathrm{N}}} + A_{bl},\ \xi^2 \underline{\underline{\mathrm{C}}}\left[\mathrm{diag}(\underline{\mathrm{N}})\right] \underline{\underline{\mathrm{C}}}^T + \sigma_{elec}^2\,\underline{\underline{1}}\Big) && (5.23)
\end{aligned}
$$

The electronic noise σ_{elec} is determined by taking the sample standard deviation of the signal before switch-on, with the baseline shape subtracted.

5.5.4. Stray Light

The real data for both systems also has some contribution from unknown reflections of the laser pulse. These each have the shape of the instrument function C and appear at approximately consistent times in each frame. Unfortunately, some of these *stray light* peaks appear on top of the TS scattering light which has very significant effect on the infered profiles in the standard analysis, especially for the Edge LIDAR system. The peaks have a largely unpredictable amplitude making them very difficult to remove by pre-processing the data and so they are modelled by adding copies of C to the post-convolution mean predicted signal $\tilde{N}$ with positions and amplitude given by an extra set of nuisance parameters. A statistical survey of the appearance of the pulses was performed on Dry-run data (where no TS light is present) and used to construct priors for the positions and amplitudes on both systems.

5.5.5. The Complete Model

Figure 5.8 shows the real data and full convolved likelihood model (equation 5.22) for two channels of the core and one of the edge LIDAR system at the MAP of an arbitrary inversion. The separate components of the model are highlighted to show that many of them are absolutely essential for the inversion to isolate the information in the TS signal. Table 5.1 gives a list of the nuisance parameters that are required to describe the model. The signal building parameters are assigned relatively weak priors and are inferred almost entirely from the data in each frame. The calibration parameters are expected to remain reasonably constant over pulses and frames and so should be assigned prior PDFs according to separate calibration experiments.

Parameter	Description
Signal building parameters: (All per channel)	
x_{bl}, y_{bl}	Baseline positioning.
x_{tp}, A_{tp}	Timing pulse position and amplitude.
x_{so}, A_{so}	Switch-on pulse position and amplitude.
N_{am}, x_{am}	Ambient light level amplitude and start position.
x_{bw}, l_{bw}	Back wall pulse position and length scale.
x_{sl}, A_{sl}, w_{Sl}	Stray light positions, amplitudes and widths.
Calibration parameters:	
Λ	Combined transmission factor.
Λ_{ch}	Relative sensitivity (per channel).
$S(\lambda)$	Normalised channel sensitivity spectrum.
ξ	Photoelectrons per bin to ADC unit factor (per channel).
Δt_{bwtp}	Timing pulse to backwall time difference (per channel).
z_{bw}	Back wall beam dump physical position along line of sight.
$V(z)$	Vignetting curve.

Table 5.1.: Nuisance Parameters involved in the full forward model.

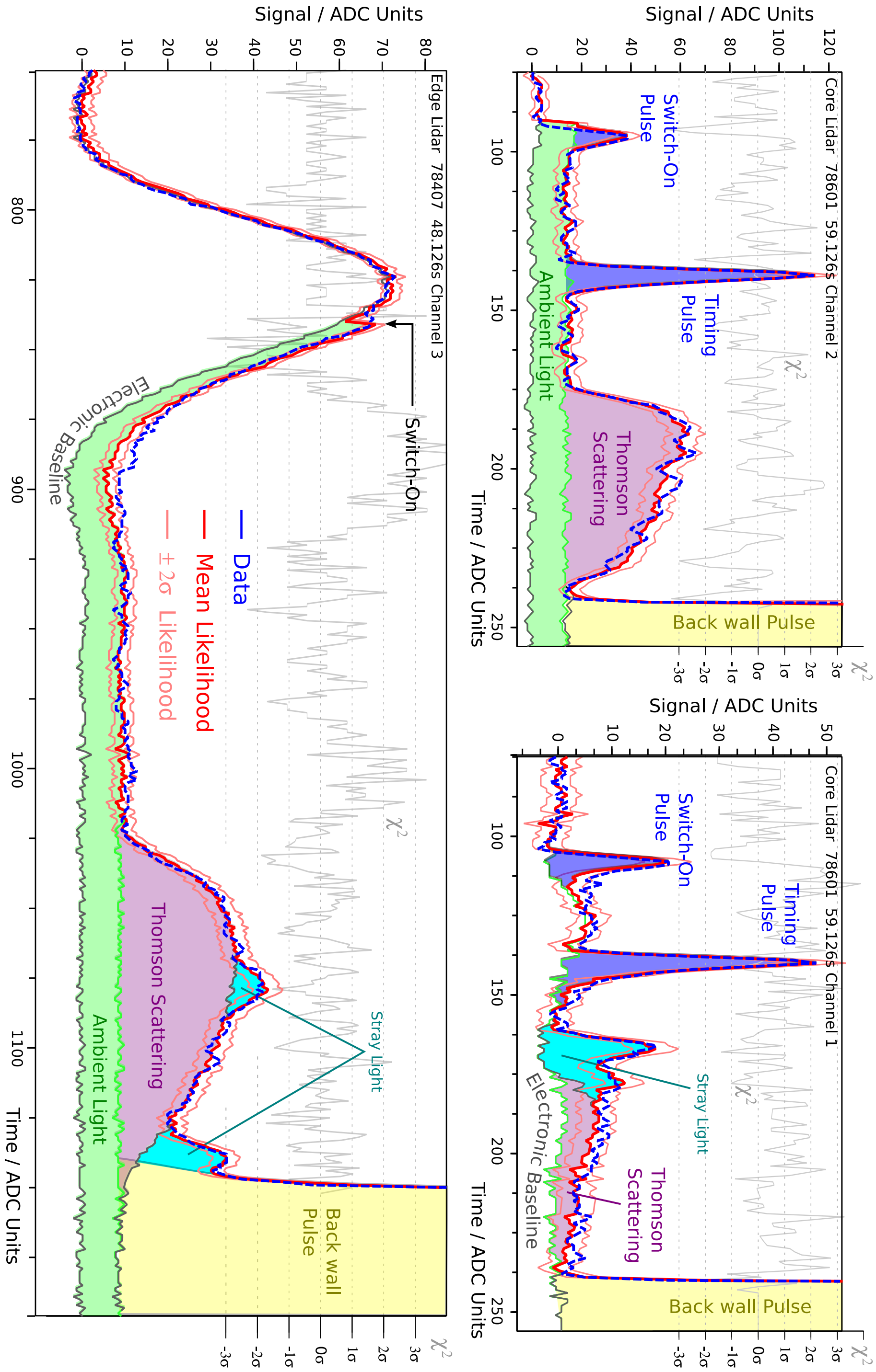

Figure 5.8.: Data (blue) and full forward model (red) for two channels of the Core LIDAR system and one from the Edge LIDAR system.

5.6. Calibrations

The LIDAR systems have an unusually large number of calibration parameters which each effect the inference of the results in different, sometimes subtle, ways. In the standard analysis, all of these are assigned fixed values determined from instrument specification or separate calibration experiments. While these values are accurate enough for that purpose, it was found that the detail of inference desired here required much greater accuracy and knowledge of uncertainty than was available. Without any ability to work with the physical equipment, a series of methods were developed to extract accurate calibration information from the vast amount of dry-run and plasma data available.

5.6.1. Instrument Functions

The convolution function C is of great importance in the analysis of both LIDAR systems. If ignored, rapid changes in the density and temperature profiles like the H-mode pedestal are not correctly inferred. The major effect on the density gradient is approximately as if the real n_e profile were itself convolved with C. The effect on T_e is more complicated as the inferred T_e at a point is biased towards the T_e of higher density plasma within the scattering volume. Figure 5.9 shows the posterior of an inversion using the simple approximation (equation 5.21) which ignores the convolution, on noise-less simulated core LIDAR data generated using the full convolved model (equation 5.22). Whether or not the sharp changes can be inferred is not important here, and the inferred MAP does not need to coincide with the original profile, but the original parameter set should be within the posterior mass (and the profile within $\pm 2\sigma$ in figure 5.9). It is clear that when not including the convolution effect, this is not the case.

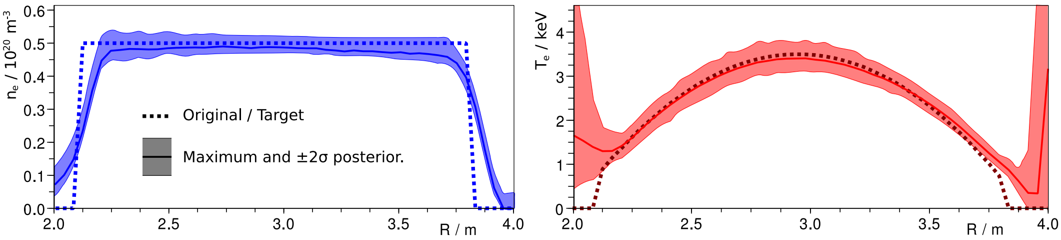

Figure 5.9.: Effect of ignoring the instrument function of the core LIDAR system. Posterior distribution without convolution in the likelihood model, based on simulated data using the full convolved forward model.

Since the standard analysis ignores the convolution, so has no need for the instrument function, an accurate record of C was not available. Fortunately, the timing pulse in the data frames should be similar to the instrument function, since it is created by feeding a small part of the laser pulse directly into the PMTs via an optic fibre. The dispersion of the fibre is assumed to be small enough not to modify the shape significantly. For a single frame of core LIDAR data, the timing pulse is covered by only around 10 ADC points and is too noisy to be of any direct use, so a procedure was developed to extract it from 130 frames of 40 dry-run pulses. Dry-run pulses were used to ensure that ambient and TS plasma light do not contaminate it.

For each frame, the timing pulse can scale in amplitude and translate in time within a known range and for the edge LIDAR system, where the laser pulse length is less stable, it may also change width. The translation is beneficial since each frame gives a glimpse of the shape at slightly different positions. Using so many frames means the common shape can be found to much higher resolution (20ps / 3mm for core LIDAR) than the ADC sample rate (200ps / 3cm). The procedure is effectively a very large fit, using a linear interpolation of a single series of knots to describe the shape and 3 parameters for each frame describing the scaling in both time and amplitude and the translation in time of the common shape onto that frame. The shape model had ~ 40 knots, spaced with approximately 4 per ADC data point. The best fit was found by iterating between the genetic algorithm and conjugate gradient searches. Figure 5.10 shows the results, which are used as the instrument function C throughout this work. Also shown is the only data initially available - a photocopy of an old photo of an oscilloscope screen from when the core LIDAR PMTs were installed in the 1980s.

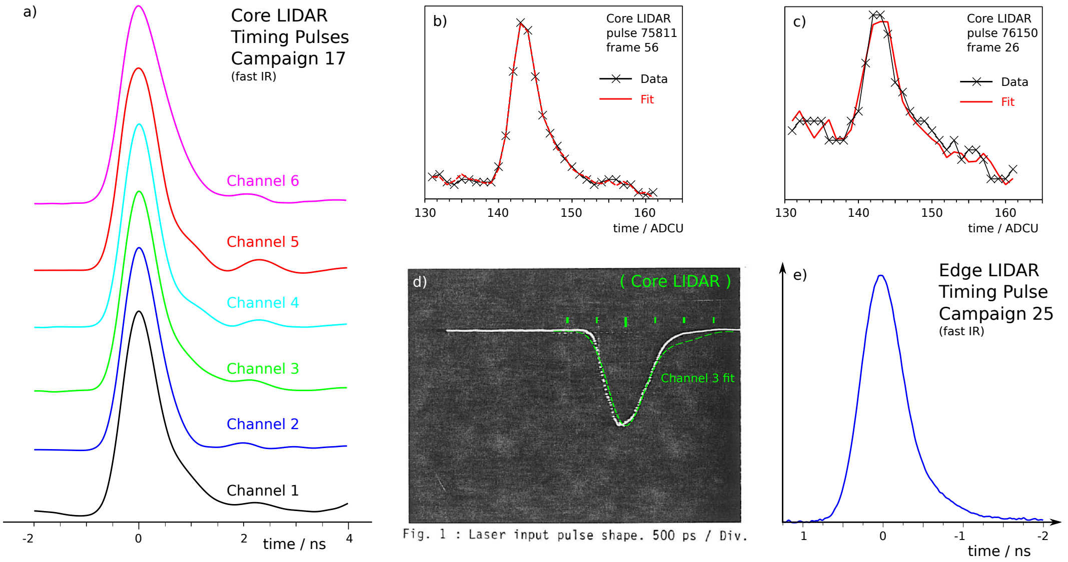

Figure 5.10.: Determination of LIDAR instrument functions using fit of common shape to the timing pulse in 130 frames of 40 dry-run pulses: a) final shape of core LIDAR timing pulses. b & c) common shape and two data frames for core LIDAR, d) Only existing information - a photo of an oscilloscope trace for core LIDAR and e) final Edge LIDAR timing pulse.

5.6.2. Photo-electron factors (ξ_{ch})

For each channel, the final ADC signal follows $D = \xi_{ch}\Lambda_*\Lambda_{ch}N_{\gamma 1}$. The photo-electron factors ξ_{ch} used in the standard analysis are calculated from the quantum efficiency and HV gain of the PMTs given in their original specifications. A large inaccuracy in these would not strongly effect the standard analysis because the relative sensitivities Λ_{ch} (see later) are actually determined from an experiment that measures the ratios of the product $\xi_{ch}\Lambda_{ch}$ between different channels, from which the assumed values of ξ_{ch} are divided out. However, the likelihood distribution variance follows $\sigma_D^2 \sim (\eta^2\xi^2\Lambda_*\Lambda_{ch}N_{\gamma 1} + \sigma_e^2)$ and so the correct values of ξ_{ch} are essential for the Bayesian analysis to correctly treat the uncertainties and to know the relative amount of trust to put in the data from each channel.

An indication that the original values were inaccurate for the core LIDAR system, and grossly incorrect for the edge LIDAR system was quickly observed by the much larger fluctuations observed in the data than predicted by the original ξs. This also highlights a way to determine these directly from the data itself. For real plasma frames, there are small regions in-between the switch-on pulse, timing pulse and TS light that show only the constant ambient light level (see figure 5.11a). The amplitude of this varies from frame to frame and is unknown in each, but the relationship of the variance with the amplitude is determined only by ξ, η and σ_e:

$$\sigma_D^2 \sim \eta^2\xi D + \sigma_e^2 \tag{5.24}$$

For 800,000 frames of 6000 pulses for core LIDAR and 17,000 frames of 3000 pulses for edge LIDAR, the sample mean $x = \langle D \rangle$ and standard deviation $y = \left\langle (D - \langle D \rangle)^2 \right\rangle$ of each of the ambient light windows were calculated. This forms a data set from which ξ can be determined, given that η is known from section 5.6.1.

A simple linear least squares fit is not sufficient because the presence of small stray light spikes in many of the frames causes distant outliers in the data. Assuming a fixed Gaussian likelihood distribution for $P(y|x)$ is also invalid because y is a sample standard deviation of the data D and so also has an uncertainty that scales with the amplitude. A linear least-squares fit also assumes a uniform prior which is not ideal because the parameter being inferred ξ, is a variance-like quantity. The likelihood distribution required is given in [71] and is in the second term of equation 5.25. The multiplier of this and the first term express the finite possibility $P_{outlier}$ of the point being an out-lier, in which case a uniform distribution represents that, once deemed an out-lier, a point is equally likely to be anywhere in a wide range $y_0 - y_1$. The 'forward function' of this fit is then just equation 5.24.

$$P(y \mid \sigma_D) = P_{outlier}\left[\frac{1}{(y_1 - y_0)}\right] + (1 - P_{outlier})\left[y^{-(n+1)} exp\left(\frac{-(n-1)\sigma_D^2}{2y^2}\right)\right] \tag{5.25}$$

Figure 5.11 shows the ambient light regions for both systems, the data set $(x,\ y)$, a kernel density plot of y for bands of x along with the prediction $\sigma_D(D)$ and its expected 2σ standard deviation at the maximum posterior value of ξ, for one channel of the core and one channel of the edge LIDAR systems.

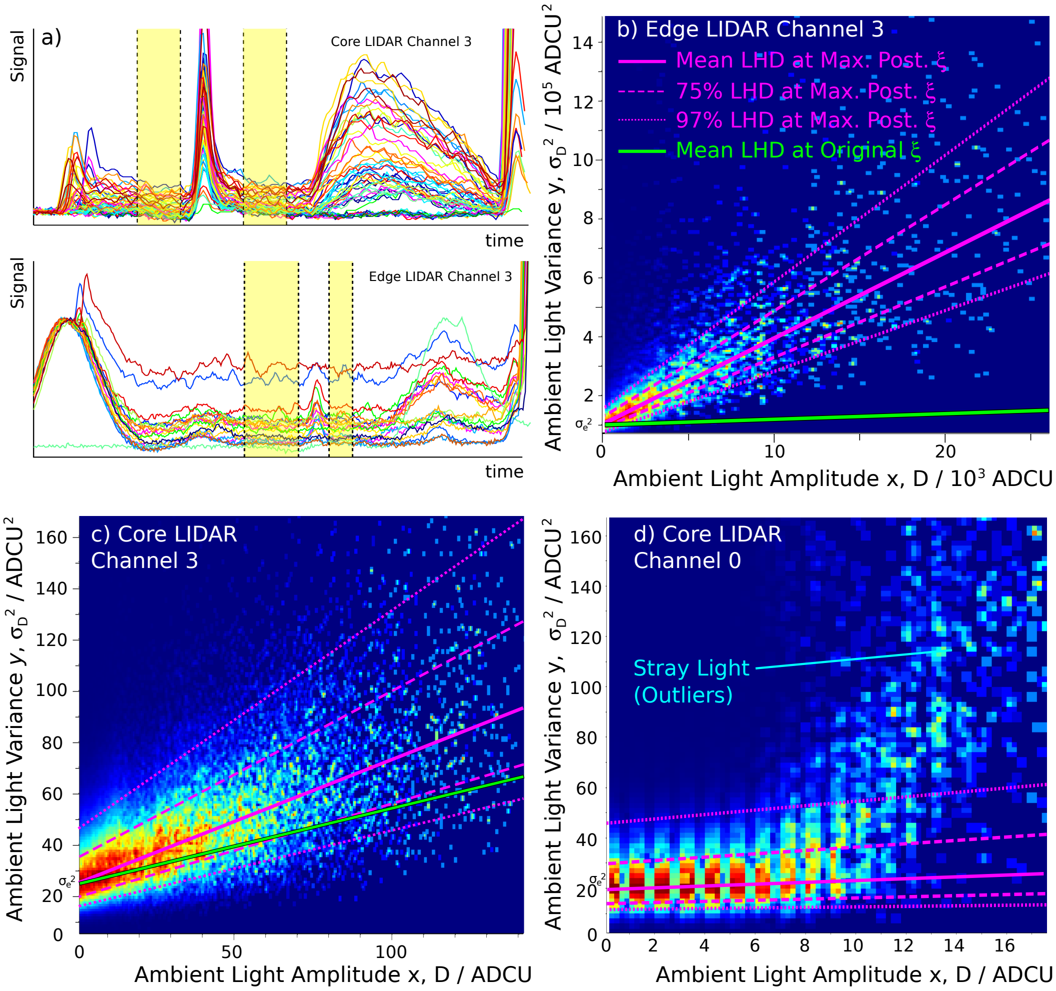

Figure 5.11.: Bayesian inference of photo-electron factors ξ from dependence of ambient light variance on amplitude. a) Regions of constant ambient level above baseline and data from some typical frames. b,c,d) Histograms in $y = \left\langle (D - \langle D \rangle)^2 \right\rangle$ for bands across $x = \langle D \rangle$ showing distribution of data in (x, y). b,c,d also show mean and spread of the likelihood (LHD - the expected distribution of y) at the MAP (best fit) and mean at the original specification values of ξ.

The procedure works well for almost all channels of both systems, with the non-outlier points in (x, y) following the second term of equation 5.25 almost exactly. The ambient light levels

System	Channel	Convolution compensation (η^2)	Original ξ	Bayesian $P(\xi \mid x, y)$
Core	1	0.121	1.18	3.63 ± 0.17
	2	0.139	1.71	3.41 ± 0.02
	3	0.122	2.59	4.20 ± 0.05
	4	0.144	1.81	2.20 ± 0.02
	5	0.127	2.94	5.31 ± 0.04
	6	0.110	1.56	3.67 ± 0.02
Edge	1	0.044	56.2	1260 ± 35
	2 ($74313 \leq$ pulse ≤ 78787)	0.044	44.6	770 ± 20
	2 ($78788 \leq$ pulse ≤ 79840)	0.044	27.0	505 ± 40
	3	0.044	25.2	690 ± 20
	4	0.044	58.0	555 ± 20

Table 5.2.: Photo-electron factors (gains) in ADCU per photo-electron per ADCU determined from specifications (original) and by Bayesian inference from ambient light variance.

on channel 1 of the core LIDAR system (bottom-right of figure 5.11) are effected strongly by the discrete levels of the ADC (1 unit) and by the fact that the ambient light regions are very small for channel 1 and the contamination by stray light very large. Table 5.2 shows the inferred values of ξ compared to the originals for both system. The reason for the discrepancies, which are well over an order of magnitude on the edge LIDAR system are not understood. It is possible that the model for the PMTs and/or ADCs is not quite correct but this will not matter now since the model with the inferred values of ξ now accurately describe the real noise levels, regardless of the reason why. It is clear that this procedure was necessary for both systems, especially the edge LIDAR system, for the Bayesian inference of n_e and T_e to properly account for the photo-electron noise.

5.6.3. Absolute Sensitivities (Λ_*)

The absolute sensitivity Λ_* is not measured for either of the LIDAR systems. For the standard analysis it is determined by a statistical comparison of interferometry data with the equivalent prediction using the standard analysis LIDAR n_e profile and fixed EFITJ flux surfaces. These inferred values for Λ_* can be used as priors here but it is usually better to apply a uniform prior and include the interferometry likelihood in the posterior so that the most consistent value for both LIDAR systems is found for the time point of interest, using the magnetic topology model of the inversion being performed.

5.6.4. Core LIDAR Relative Sensitivities (Λ_{ch})

The LIDAR system relative sensitivities Λ_{ch} principally effect the inference of the absolute T_e and are determined from separate calibration experiments by measuring $\xi_{ch}\Lambda_{ch}$ of each channel relative to the others. For the core LIDAR system, it is believed that the calibration values are accurate since the T_e results of the standard analysis have always matched those of the independently calibrated ECE diagnostic reasonably well.

Some supporting evidence can be seen from the ambient light data of section 5.6.2. Since the spectral ranges of the core LIDAR channels are fairly wide, it is reasonable to suspect that the majority of the ambient light might come from Bremsstrahlung emission since any line emission is usually very narrow and even if large in amplitude will not give a large overall power/intensity. The Bremsstrahlung emission photon count $N(\lambda)$ approximately follows $1/\lambda$ over the LIDAR spectrometer range so the ratio of the ambient light levels between two channels should be:

$$\frac{\langle D_i \rangle}{\langle D_j \rangle} = \frac{\xi_i \Lambda_i \int \lambda^{-1} S_i(\lambda) d\lambda}{\xi_j \Lambda_j \int \lambda^{-1} S_j(\lambda) d\lambda} \tag{5.26}$$

A major line emission that might have a large effect is the D_α line, the intensity of which can be found from the visible spectroscopy diagnostic. Figure 5.12 shows $\langle D_i \rangle$ vs $\langle D_j \rangle$ for each combination of channels coloured by D_α emission intensity from visible spectroscopy. Channel 1 is not included as the ambient levels are too low to be useful for this exercise. The expected ratio according to the calibration values of Λ_{ch} are shown and sit in many cases approximately along the bulk of the data points. The D_α emission line sits inside channel 2's spectral range (see figure 5.1) and can only give a positive contribution to $\langle D_2 \rangle$. The first four plots of 5.12 involve channel 2, and show that the points are higher in $\langle D_2 \rangle$ than the expected ratio by an amount that clearly depends on the D_α emission. Due to this, it is difficult to tell if Λ_2 is accurate. The remaining plots suggest that Λ_3, Λ_5 and Λ_6 are all very accurate but that there might be a small inaccuracy in Λ_4.

This investigation was carried out here only to determine the level of confidence in the calibration Λ_{ch} values. Based on this and the general agreement of core LIDAR with ECE, a very narrow prior can be used or they can be fixed at their calibration values since the uncertainty is unlikely to contribute a major part of the uncertainty in n_e and T_e. There is also a great deal of scope for further investigation of the technique. With a proper spectroscopic survey and/or using more detailed data from the visible spectroscopy any other spectroscopic diagnostics it might be possible to automatically calibrate LIDAR systems using this technique alone. A major advantage is that it includes the full optical path of the diagnostic as it is used, which is not true for the calibration experiment at JET because part the collection windows (see figure 5.3) must be removed to gain access to the vessel. It also would then be relatively easy to develop a system which automatically adapts to any changes in the collection optics

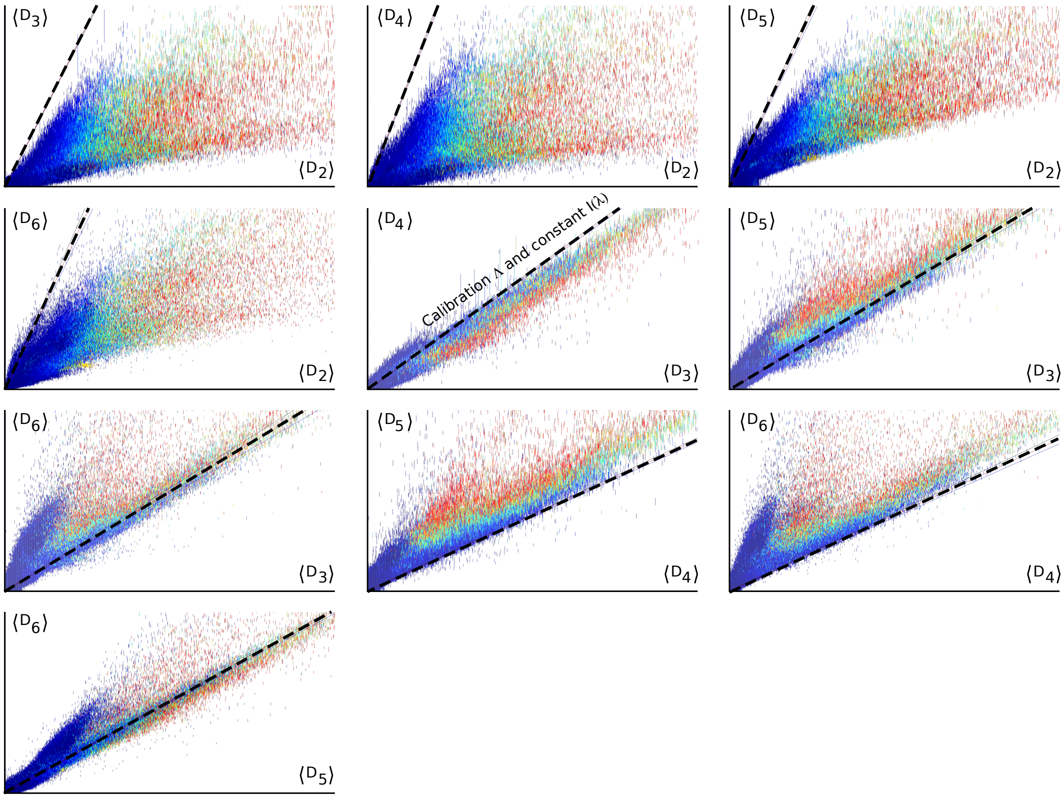

Figure 5.12.: Ratios of average ambient light intensity for channels 2-6 of core LIDAR (small coloured lines) compared to anticipated ratios using calibration values of channel relative sensitivities Λ_{ch} and assuming $N(\lambda) \propto 1/\lambda$ spectrum (broken lines). Points are coloured by D_α line emission intensity, which sits within channel 2's spectral range.

without requiring access to any of the hardware - a particularly useful feature for ITER and reactor devices where the high particle fluxes will degrade the optics over time and the high radiation levels will make manipulation of in-vessel hardware difficult.

5.6.5. Edge LIDAR Relative Sensitivities (Λ_{ch})

Unlike with the core LIDAR system, the agreement between the standard analysis of edge LIDAR and the other T_e diagnostics is relatively poor despite the use of exactly the same calibration procedure for both LIDAR systems. The large uncertainty in the positional calibration of edge LIDAR, along with difficulties in relating it to the other diagnostics due to the uncertainty in the flux surfaces (and assumptions of constancy of T_e/n_e on them) makes it difficult to quantify this or to be sure that inaccurate Λ_{ch}s are the cause. In fact, the common interpretation so far, has been that the temperatures are correct and that the edge LIDAR line of sight penetrates less deeply into the plasma than the EFITJ flux surfaces suggest, giving the temperatures at the pedestal base. The full integrated approach shows that it is almost certain that the relative sensitivities Λ_{ch} for the edge LIDAR system are significantly incorrect (by a factor ≈ 2.3). Using the standard analysis techniques, this had not been seen and it would be very difficult to correctly handle it at the same time as the other calibration and noise uncertainties.

The ambient light ratios procedure used for the core LIDAR system is unlikely to be applicable to the edge LIDAR system because the optics view the edge and divertor regions where the plasma is cold and contribution from spectral lines above the Bremsstrahlung is likely to be large. Figure 5.13 shows the ambient light data that is observed and the expected ratios if the spectrum were still approximately $N(\lambda) \propto 1/\lambda$, for the original calibration and for the widest possible range of values that give temperatures consistent with the core LIDAR (these are found later, in section 5.7.3).

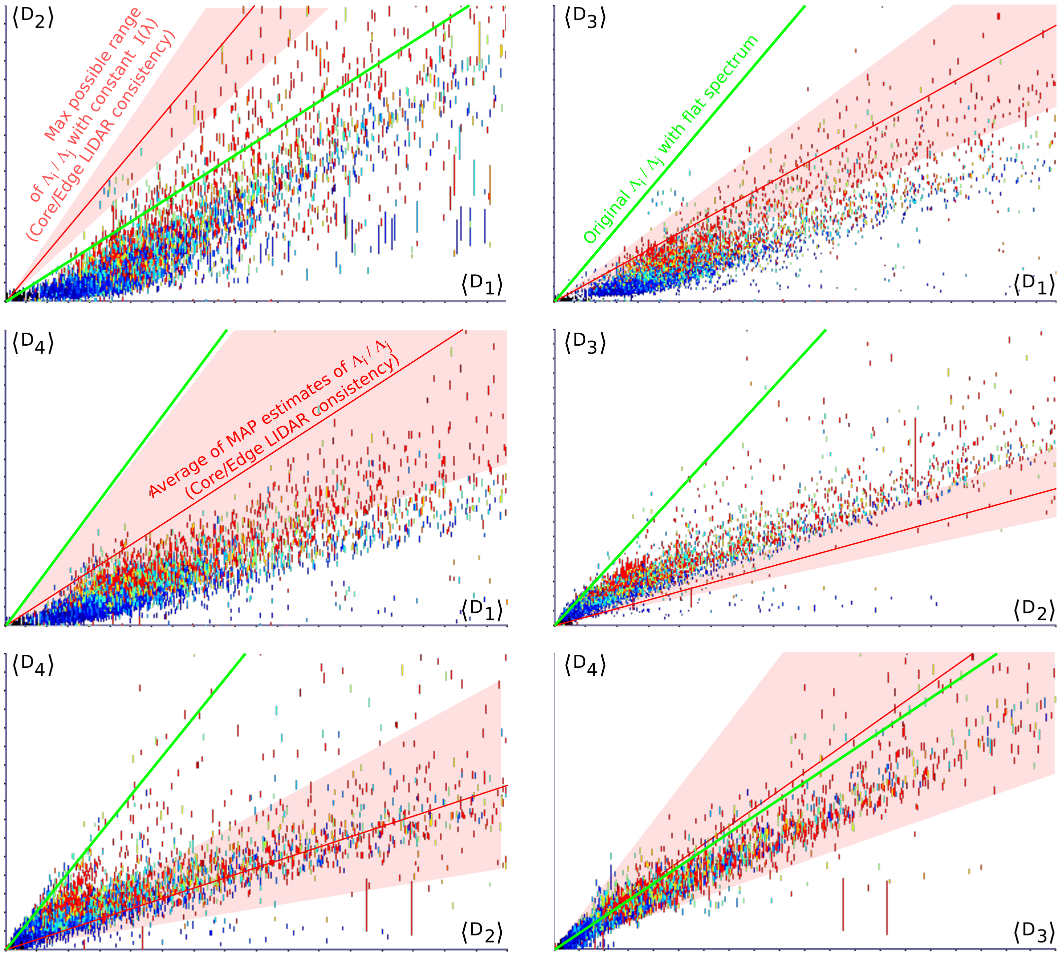

Figure 5.13.: Ratios of average ambient light intensity for all 4 channels of edge LIDAR (coloured lines) compared to anticipated ratios using calibration values of channel relative sensitivities Λ_{ch} and assuming $1/\lambda$ spectrum for $N(\lambda)$ (green solid lines). Also shown is the expected ratios from the widest range of possible Λ_{ch} values obtained in section 5.7.3 (pink band).

While it does not seem to be consistent with the possible relative sensitivities and the simple uniform spectrum assumption, there is a strong correlation in the observed ambient levels of the different channels despite the very wide range of plasmas (fuelling, I_p, B_t, n_e, T_e etc). This suggests that a proper study of the visible spectrum in edge LIDARs viewing area and modelling of the expected ratios of integrated intensity in each channel could yield an accurate calibration. Without this, and having cast significant doubt on the original calibration values, the relative sensitivities Λ_{ch} are initially assumed almost completely unknown and $P(\Lambda_{ch})$ assigned a uniform or very wide prior. In section 5.7.3, more accurate values for Λ_{ch} are found from posteriors of full inversions including the core LIDAR likelihood.

5.6.6. Timings and position calibrations

Various timing and position parameters are present in the model. Each system has a global timing parameter across all channels which effects the position of the inferred n_e and T_e profiles along the system's line of sight. If assigned a uniform prior, the global timing is usually easily inferred from the data because the core system sees both the inboard and outboard sides of the plasma. The parametrisation of the profiles as constant on flux surfaces results in an effective automatic centering on the consistent solution. The edge LIDAR system is then effectively aligned to this.

A set of parameters are used to describe the difference in timing between the channels and these can have a small effect on the shape of both the T_e and n_e profiles. For core LIDAR, the typical gradual monotonic increase then decrease in T_e across the core plasma makes them easy to infer even if assigned a uniform prior. The conceptually simplified explanation is that the signal must rise across the channels in order then fall in the opposite order.

The same is not true of the edge LIDAR system, where T_e along the line of sight rises to near its maximum and does not fall much before the signal is lost beneath the back-wall light. Uncertainty in the timing differences would have a large effect on the T_e pedestal gradients and widths inferred. By design, there should be no timing difference because, unlike core LIDAR, a single common-trigger ADC is used and the system is designed so that all optical paths and cables are the same length. The edge LIDAR channel to channel timing differences are therefore assumed to be zero for this work. Individual large stray-light peaks that occasionally appear in multiple channels also support this, since they are rarely located more than 1 data point apart in different channels. This suggests the inter-channel timing misalignment is at most $7.5mm$ (1 ADC period).

5.7. Inversion

5.7.1. Free shape, free calibration inference

With all of the nuisance and calibration parameters assigned appropriate priors and the desired model for $n_e(R, Z)$ and $T_e(R, Z)$ connected, inversions can be performed with both LIDAR systems and the interferometry. The complete combination can be used for time points at which the data is captured sufficiently close together that the magnetic geometry, n_e, and T_e will not have changed significantly, so that all three systems are looking at approximately the same plasma. The main limitation is the 6 time points at 1s intervals of edge LIDAR, since core LIDAR will usually have a time point within 5ms of these and the interferometry within 1ms.

With so many of the calibration and nuisance parameters assigned weak priors, the posterior is very broad (i.e there is large uncertainty) when the LIDAR signal levels are low. For high density and hence high TS light level plasmas, there is sufficient information to infer n_e and T_e reasonably accurately despite the lack of calibration information. Initially, fixed EFITJ flux surfaces are used and figure 5.14 shows the marginal profiles of the posterior $P(n_e, T_e, \mu_{calibs} \mid D_{interf}, D^{core}_{LIDAR}, D^{edge}_{LIDAR}, j^{EFITJ}_{\phi})$, for an H-mode plasma shortly before a Type-I ELM. The n_e and T_e profiles are parameterised with the linear-interpolation of 50 knots with greater concentration at the edge. The profile priors assigned are second-differential smoothing priors relatively strong in the core and weak in the edge region.

For this fairly recent pulse, the High Resolution Thomson Scattering (HRTS) system is available and is shown in figure 5.14 for comparison, as the HRTS data is not used in the inversion in any way. The results of the standard analysis of the same data on which the inversion is performed are also shown. It is immediately clear that a far more accurate result is obtained than by analysing each diagnostic independently.

The overall magnitude of the density differs slightly to the standard analysis of core LIDAR due to the inclusion of the interferometry diagnostic's accurate integral information. The pedestal region shows the most improvement, where the combined analysis agrees very well with the HRTS when the standard results show little resemblance. The core LIDAR stand-alone results show no T_e pedestal at all and only a small slow drop in n_e, which is due to the effective 12cm resolution that comes from the convolution effect which the standard analysis ignores. It seems logical to assume that the pedestal information must be derived almost entirely from the edge LIDAR data but the core LIDAR also provides some critical parts of the information. Edge LIDAR does not provide the absolute magnitude of either n_e or T_e because of the lack

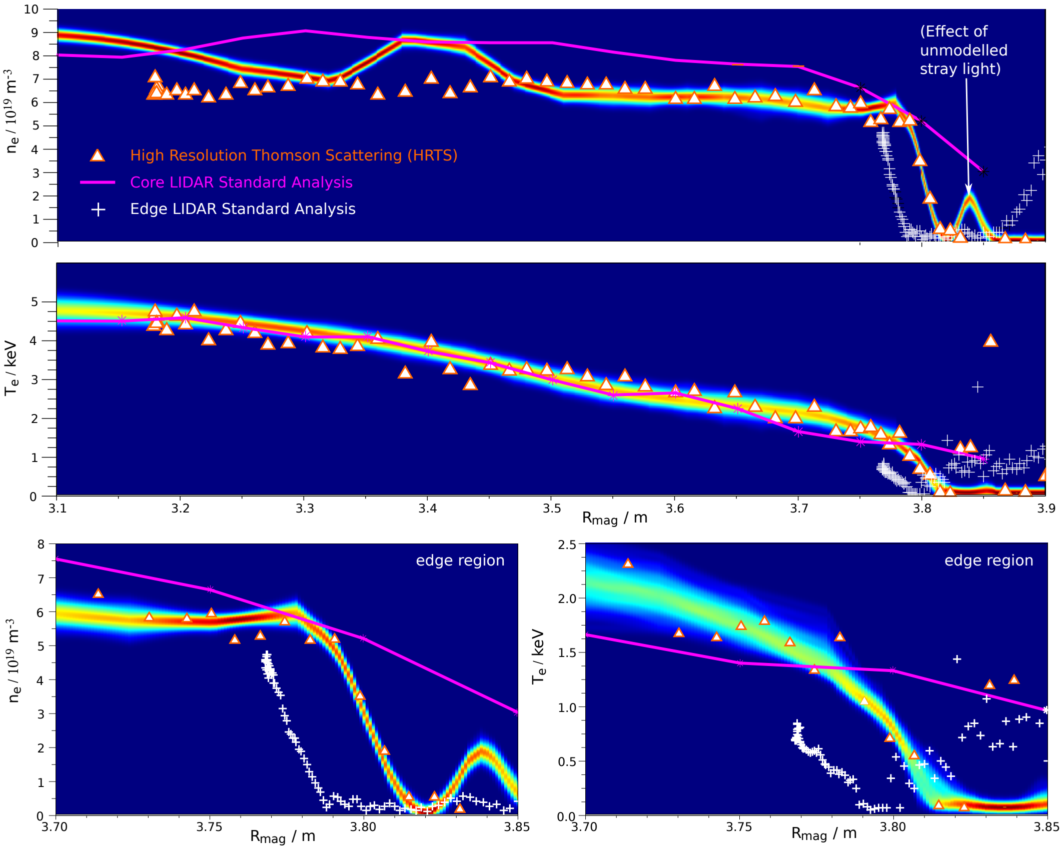

Figure 5.14.: Profile marginals (colour banding) of posterior $P(n_e,\ T_e,\ \mu_{calibs} \mid D_{interf},\ D^{core}_{LIDAR},\ D^{edge}_{LIDAR},\ j_\phi^{EFITJ})$. Inference of free-form n_e and T_e profiles from combination of interferometry, core LIDAR and edge LIDAR based on fixed flux surfaces from EFITJ. Also shown are the standard analysis of D^{core}_{LIDAR} and D^{edge}_{LIDAR} and of the HRTS diagnostic. The plasma is a particularly high-density but otherwise typical Type-I ELMy H-mode with each diagnostic data sampled within 1ms of each other, $\sim 10ms$ before an ELM.

of knowledge of the spectrometer sensitivity parameters Λ_* and Λ_{ch}, although it does provide accurate information about their shape. While the core LIDAR system cannot see the pedestal shape, it effectively provides something similar to the integral of T_e in the last 12cm. From both pieces of information the full T_e profile is inferred, along with an estimate of the edge LIDAR Λ_{ch} values. The situation for n_e is slightly more complex, since core LIDAR also has uncertain Λ_* and cannot know the n_e magnitude. It does however couple the edge n_e 12cm integral to the core profile, which gets some shape information from core LIDAR and magnitude information from the interferometry. In a similar way, all of the timing and positional parameters which were assumed unknown and assigned weak priors are also inferred by self-consistency.

While it is an interesting exercise to think about from where the information is effectively coming, the power of the Bayesian analysis of the combined system is that none of this needs to be even considered. The profiles are obtained by carefully expressing what is known a-priori and then asking what it implies mathematically can be known about n_e and T_e. In many cases, especially some with low signal to noise, it is surprising how much information is really present in the data, when the standard analysis produces results that many scientists consider too noisy and inaccurate to even consider working with. The inferred profiles usually match the HRTS results better than can be achieved by arbitrarily shifting and scaling the stand-alone analysis profiles to agree, a method that is often used to resolve the inconsistencies, but rarely carefully justified.

Flux surface uncertainties

Because the fixed EFITJ flux surfaces are used, the n_e profile in figure 5.14 shows similar artifacts in the core as those seen in the EFITJ based interferometry-only n_e inversions in section 4.1.2. To account for the flux surface uncertainties, the current tomography model can be used for j_ϕ in the combined inversion, together with the CAR prior. The posterior is then $P(n_e, T_e, j_\phi, \mu_{calibs} \mid D_{mags}, D_{interf}, D^{core}_{LIDAR}, D^{edge}_{LIDAR})$, which is shown in figure 5.15. The inversion uses the same parameters and priors so, apart from including the flux surface uncertainty, it is identical to that used to obtain figure 5.14.

As expected, the uncertainty is now much larger showing that the uncertainty in the flux surfaces makes a large contribution to that in the profiles. The n_e pedestal position is particularly uncertain, although the shape does not vary significantly (seen in the individual samples shown for the edge region). The T_e pedestal does show that a range of shapes are possible. Some show a sharp transition between the core gradient and the ETB while others have a slowly increasing gradient. This variation in shape is seen for many pulses, in both R_{mag} or ψ_N space. It comes from the fact that the flux surface uncertainty given only the magnetic measurements is sufficiently large, that a wide range of pedestal shapes can be made to match the shape of the edge LIDAR data by appropriate manipulation of the flux surfaces along its line of sight. Figure 5.16 shows the flux surfaces from the full posterior compared to those of the fixed EFITJ solution. The large differences between the absolute values of ψ_N are not important, since the profiles are all positioned by self consistency and are not in any way tied to the LCFS. The shape is however important, since the assumption of constancy of n_e and T_e on the flux surfaces requires the edge LIDAR data of the inboard edge, to match that of the outboard. Consistency of this assumption with the data requires the plasma to be narrower just above the X-point than in the EFITJ solution.

As with the polarimetry investigations, the choice must be made between the large but

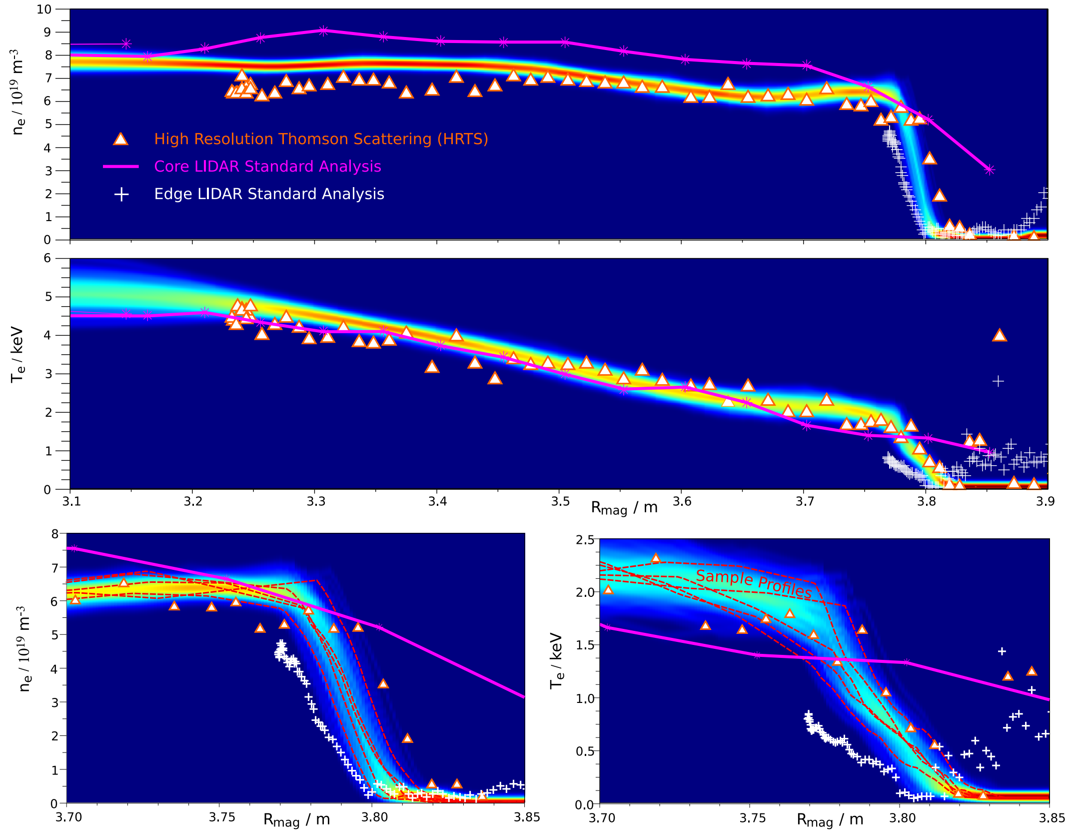

Figure 5.15.: Profile marginals (colour banding) of posterior $P(n_e,\ T_e,\ j_\phi,\ \mu_{calibs}\ |\ D_{mags},\ D_{interf},\ D^{core}_{LIDAR},\ D^{edge}_{LIDAR})$. Inference of full free-form n_e and T_e profiles from combination of magnetics coils, interferometry, core LIDAR and edge LIDAR including uncertainty due to flux surface uncertainty. Also shown are the standard analysis of D^{core}_{LIDAR} and D^{edge}_{LIDAR} and of the HRTS diagnostic and several individual samples of the posterior (red).

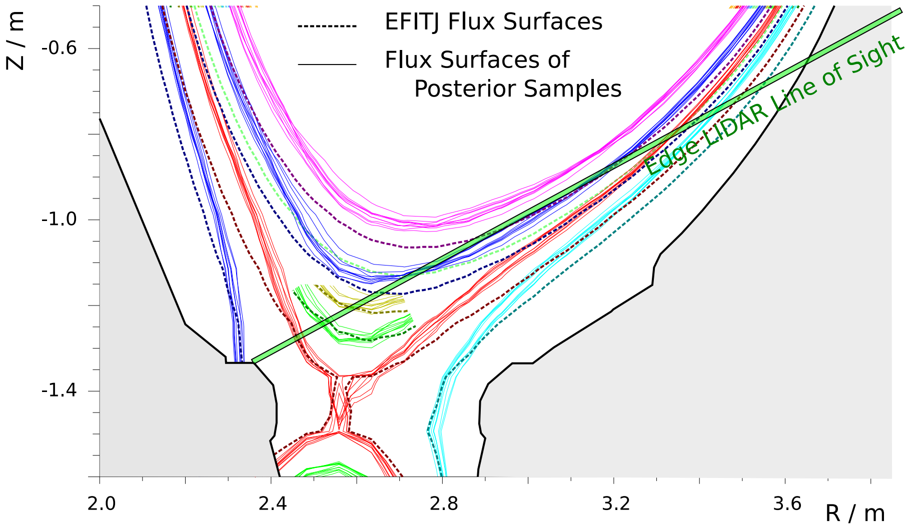

Figure 5.16.: Flux surfaces from samples (solid) of posterior $P(n_e,\ T_e,\ j_\phi,\ \mu_{calibs}\ |\ D_{mags},\ D_{interf},\ D^{core}_{LIDAR},\ D^{edge}_{LIDAR})$ compared to EFITJ solution (dashed) and edge LIDAR line of sight.

quantified uncertainty of the combined inversion or the systematic uncertainty in the EFITJ surfaces that is smaller but cannot be accounted for. This recurring problem is revisited in chapter 6 but for the remainder of this chapter, the EFITJ flux surfaces are used due to the practical issue of inversion speed. The large collection of very different calibration parameters and their interplay makes the core/edge LIDAR combinations produce very non-linear, multi-modal and extremely complex posteriors. The combination of this with the high dimensionality of j_ϕ takes an impractically long time to explore for repeated use.

5.7.2. Pedestal Parameterisations

For low density and hence high noise pulses/time-points, the pedestal shape inferred becomes very uncertain. It is usually still possible to extract broader information about the pedestals by using a stronger parameterisation with fewer degrees of freedom. Particular quantities of interest are the pedestal height ($n_{e,ped}$, $T_{e,ped}$), width (Δ) and hence gradient, and pedestal centre position ($\psi_{N0} = \psi_N$ at half the pedestal height). A common choice used to parameterise the pedestal is the modified hyperbolic tangent (mtanh)[72]. This was proposed as a model for n_e, based on the assumption that the formation of the n_e pedestal is largely due to the high charged particle source rate at the edge, due to the neutral density falling rapidly inside of the separatrix [73]. This is itself supported by the presence of the n_e pedestal in L-mode when the ETB does not exist (see for example, the interferometry inversion in section 4.1.3, figure 4.3). The inferred shapes for n_e in figures 5.14 and 5.15 also support this. For T_e, where the pedestal is formed by the suppression of transport in the ETB, it is not clear that fitting a mtanh function is justified.

Figure 5.17 shows a schematic of the two parameterisations and some n_e profile samples of an mtanh parameterised inversion for a lower density plasma. The low signal level can be seen to dramatically effect the core LIDAR standard analysis where as with the strong parameterisation, the Bayesian inversion is able to recover the profile well.

5.7.3. Inference of calibration parameters

Whether using a free-form profile or the mtanh/linear pedestal parameterisations, each posterior also contains the inferred calibration parameters. For example, the marginal posterior $P(\Lambda_{ch} \,|\, D_{mags}, D_{interf}, D_{LIDARs})$ gives the range of possible sets of values for the edge LIDAR relative sensitivities that are at all consistent with all the diagnostic data, parameterisations, priors and flux surface constancy. Figure 5.18 shows the relative sensitivity ratio Λ_3/Λ_2 at the

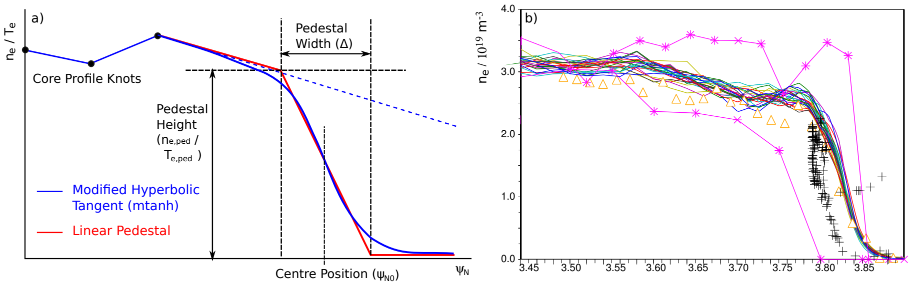

Figure 5.17.: a) Schematic of modified hyperbolic tangent (mtanh) and linear pedestal parameterisations, coupled to the normal linear-interpolated knots for the core. b) Samples for n_e of an inversion using the mtanh parameterisation on a plasma that gives low TS signal levels.

MAP estimate of $P(n_e, T_e, \mu_{calibs} \,|\, D_{interf}, D_{LIDARs}, j_{\phi}^{EFIT})$ for 1500 time-points over several campaigns using the linear-pedestal parameterisation. For a few of these, the 2σ width of the full posteriors using fixed (EFITJ) or free (CT) flux surfaces are shown to give an idea of the expected spread (The non-linear exploration is too computationally expensive to be performed for every point).

For comparison, figure 5.19 shows MAP estimates of $P(\Lambda_{ch},\, \mu_{calibs} \,|\, D_{LIDAR}^{edge},\, n_e^{HRTS},\, T_e^{HRTS},\, j_{\phi}^{EFIT})$ - the ratios Λ_3/Λ_2 required to agree with the HRTS standard analysis n_e/T_e profiles. These are only shown to give support to the ratios inferred from the combined analysis because, while they appear cleaner, they cannot be used since the HRTS has not been thoroughly modelled and the effect of its own calibration uncertainties have not been investigated here. For the pulses shown, the JET HRTS system was itself cross-calibrated from the ECE diagnostic [74].

If it is assumed that the relative optical sensitivity of the equipment will not change rapidly between pulses or time-points, the uncertainties can be reduced by taking the moving average of the MAP estimates. These are shown in figure 5.18/5.19 along with the $\pm 2\sigma$ spread of the MAPs. From this, a slow drift can be seen in Λ_3/Λ_2, which is probably due to deterioration of the diagnostic optics over the campaign.

For future inversions, the priors for the edge LIDAR Λ_{ch}s can be assigned distributions covering the range of values seen in figure 5.18a which will improve the inference of the pedestal T_e. It is also possible to use these values in the standard analysis which brings the T_e magnitude in agreement with the other diagnostics. Strictly, this prior can only be used on inversions of data that was not included in acquiring the prior but the information will be sufficiently diluted by other pulses to make little difference to the result.

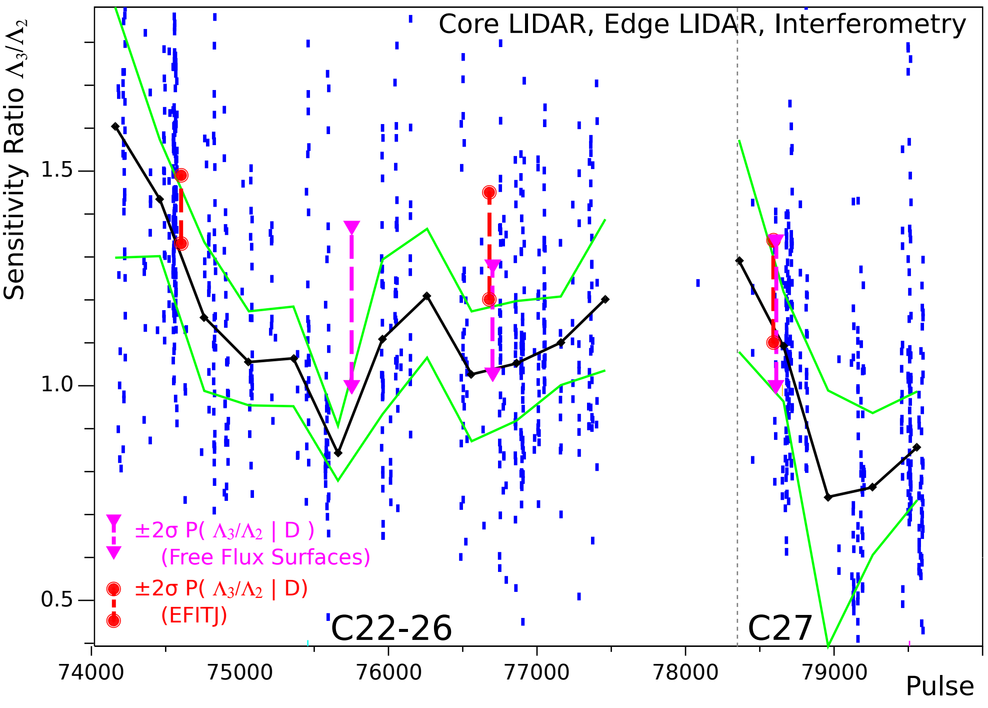

Figure 5.18.: Inferred sensitivity ratios Λ_3/Λ_2 from the combined analysis of core LIDAR, edge LIDAR and interferometry. The original calibration value, used by the standard analysis was 2.8 (off the scale).

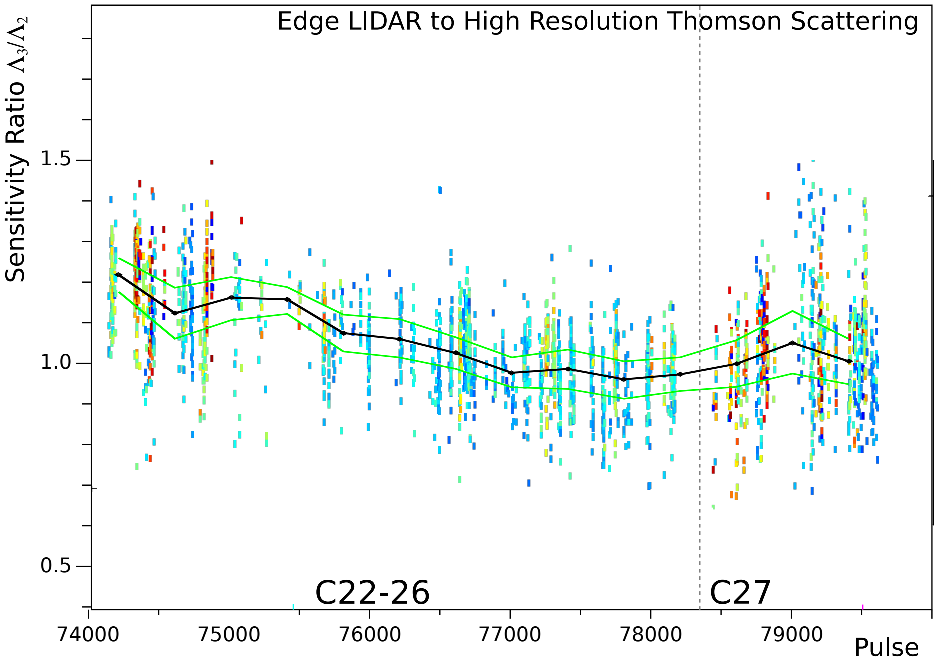

Figure 5.19.: Inferred sensitivity ratios Λ_3/Λ_2 by direct fitting to HRTS.

5.8. Pedestal Evolution

With the calibration values inferred, the combined inversion benefits from the very high spatial resolution information that the edge LIDAR system can give, for plasmas with a reasonably high TS light level (high density and high incident laser intensity). Such plasmas allow the inference of very accurate information about the H-mode pedestal. For a single time point, this is more accurate and has a higher resolution than even the new HRTS system, which gives at most around 10 points in the pedestal region, usually many less. Studies of the pedestal dynamics have previously been conducted at JET using the HRTS diagnostic [75] by averaging profiles over many time points and in some cases [76] moving the plasma by a few cm, in order to capture the required resolution. This averaging was done over many time points in different parts of the shot, at roughly the same position in the inter-ELM cycle, under the assumption that the plasma will have the same n_e and T_e profiles, at a given point in every ELM cycle. The edge LIDAR system provides a much lower noise and higher resolution snapshot of the pedestal at each time point, but unfortunately, it only records data at 6 time points separated by 1s intervals in each pulse. These fall effectively randomly within the ELM cycle.

In order to examine the inter-ELM evolution, 6 pulses were found that have a high electron density, high edge LIDAR laser intensity and almost identical global plasma configuration. The global parameters over the main heating phase of the pulses are given in table 5.3.

Pulses	78596 - 78601		
Plasma Current I_p	1.7 MA		
Vacuum Toroidal Field B_t	2.4 T		
Core n_e, T_e	$\approx 8 \times 10^{19} m^{-3}$,	$5keV$	
Heating Power	$13MW$ NBI,	$1MW$ ICRH,	$< 1MA$ Ohmic. (for 8s)
Inter-ELM period	$\approx 200ms$		

Table 5.3.: Main parameters for series of pulses analysed.

The pulses which fit the requirements were part of the ELM pellet pacing experiments[77], where small frozen deuterium pellets are fired into the plasma edge at high frequency, with the intention of deliberately triggering ELMs. Given this intention, it is likely that the pellets will effect the pedestal immediately after injection. However, in these plasmas, the pellets are injected at $20Hz$ and consist of a very small amount of fuel, compared to the high density of the plasma. The fraction of these that successfully reach the plasma is also low, with almost none triggering ELMs for most of the pulses. No lasting effect of the pellets is seen on the high-speed interferometry measurements[78]. It is therefore unlikely that at a randomly selected time (e.g an edge LIDAR time point), the plasma will be significantly different to an identical pulse without pellet injection.

5.8.1. Inversion details and MAP results

To analyse the pedestal shape, the full free-form profile inversion of section 5.7.1 was performed on the 28 time points in the main heating phase of the pulses. Second differential smoothing priors were assigned for both profiles with stronger smoothing in the region covered only by core LIDAR and weaker in region covered by edge LIDAR. The strengths were selected to be just strong enough to suppress extreme local noise in both systems, without reducing the inferred gradients significantly. Table 5.4 gives the smoothing strength σ, and the gradient below which it will be easily overridden by the data (2σ). Figure 5.20 shows the MAP results for all time points, coloured according to the time between the preceding ELM and the time point of the inversion Δt. There remains some stray-light contributions for which a satisfactory model was not developed, which contaminates the inferred profiles at $\psi_N \approx 1.03, R_{Mag} \approx 3.85$ so this part of all of the profiles is shown with reduced intensity.

Profile	n_e	T_e
Core $\sigma_{d^2f/d\psi_N^2}$	$10^{21}m^{-3}$	$50keV$
Edge $\sigma_{d^2f/d\psi_N^2}$	$10^{22}m^{-3}$	$500keV$
Max gradient (core)	$10^{19}m^{-3}$ per cm	$1keV$ per cm
Max gradient (edge)	$2 \times 10^{20}m^{-3}$ per cm	$8keV$ per cm

Table 5.4.: Second differential smoothing priors used for pedestal evolution inversions.

5.8.2. Early ELM-cycle profiles

The effect of the ELM-crash is immediately evident from the early profiles (blue) of figure 5.20. The height of the temperature pedestal drops and the earliest profile ($\Delta t = 3ms$, shown as a broken line) shows a return almost to the linear, L-mode like shape for T_e. The corresponding density profile also shows a significantly reduced gradient, but with the density reducing for $\psi_N < 0.97$ and increasing for $\psi_N > 0.97$, unlike T_e, in which all parts of the profiles are at lower T_e after the ELM. This was also seen in the HRTS investigation [75], which showed a clear pivoting of the density pedestal gradient around the pedestal centre (defined as the point where $n_e \approx \frac{1}{2}n_{e,ped}$, at $\psi_N \approx 0.97$ in this case). The MAP profiles here do not disagree with this but only the 1^{st} profile ($\Delta t = 3ms$) and the 4^{th} ($\Delta t = 9ms$), show a significantly different profile shape and height to the rest, suggesting that the density profile returns rapidly (by $\Delta t \approx 10ms$ at the latest) to its pre-ELM form. The T_e pedestal rises more slowly, reaching a height and shape similar to the late cycle profiles after $\Delta t \approx 40ms$, as can be seen from figure 5.21, which shows the T_e pedestal profiles in ψ_N grouped into sets of 5.

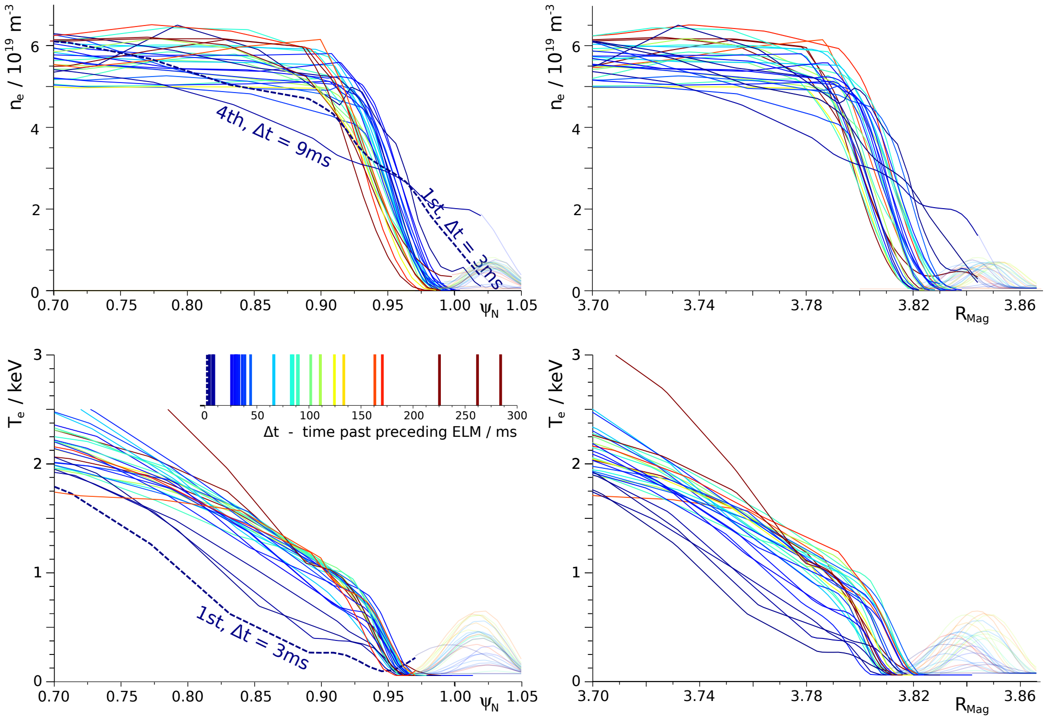

Figure 5.20.: MAP n_e (top) and T_e (bottom) profiles for 28 edge LIDAR time points in 6 similar pulses, coloured by fraction of time point through ELM cycle, with blue earlier and red later. Parts of profile with low intensity are polluted by unmodelled stray light. Profiles are plotted in both normalised flux (ψ_N - left) and magnetic mid-plane real space (R_{Mag} - right).

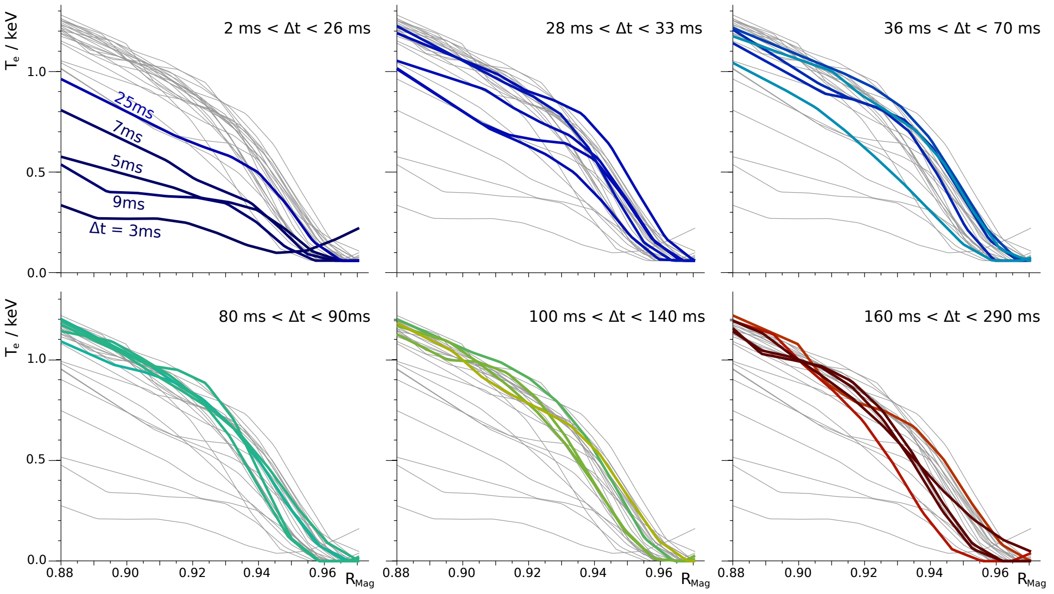

Figure 5.21.: MAP T_e pedestal profiles, grouped into sets of ~ 5, spanning successive ranges of time after preceeding ELM Δt. All 28 profiles are shown in every panel in grey, to aid comparison.

5.8.3. Pedestal position

An interesting feature of figures 5.20 and 5.21 is that the T_e pedestal is stationary in ψ_N, while it appears to have no fixed position in R_{mag}. The inversion assigns no importance to any particular value of ψ_N and the profile positions are inferred entirely from the data and flux surface consistency, so it is unlikely that this is an artifact of the inversion set-up. Considering the inferred pedestal position in real space in terms of R_{mag}, could be unecessarily introducing further errors from the EFITJ flux surfaces. The lowest noise data, and hence strongest constraint on the inferred profiles, comes from the outboard pedestal on the edge LIDAR line of sight, some distance from the outboard midplane. It is more useful to directly asses the pedestal position as a coordinate along the edge LIDAR line of sight. To examine the variations of the pedestal position in both flux and real space in more detail, figure 5.22 shows the position of the EFITJ $\psi_N = 0.91$ (arbitrary selection) contour along the edge LIDAR line of sight in black triangles for each of the 28 time points, against Δt. A 10 point moving average is also plotted for this and for intervals of 0.01 in ψ_N, which shows the general trend of the separatrix moving outwards as the plasma pressure grows over the ELM cycle. The real space positions along the line of sight, of $n_e = 3.0 \times 10^{19} m^{-3}$ and $T_e = 400eV$ are also plotted, to show the movement of the pedestals. These values of n_e and T_e are approximately half of the average pedestal height, adjusted to ensure a good estimate of the pedestal position for as many of the profiles as possible. For most of the ELM cycle ($\Delta t > 30ms$), the T_e pedestal position can been seen to moves outwards, following the movement of the separatrix, while the density pedestal shows no clear trend, possibly remaining fixed to a real space position.

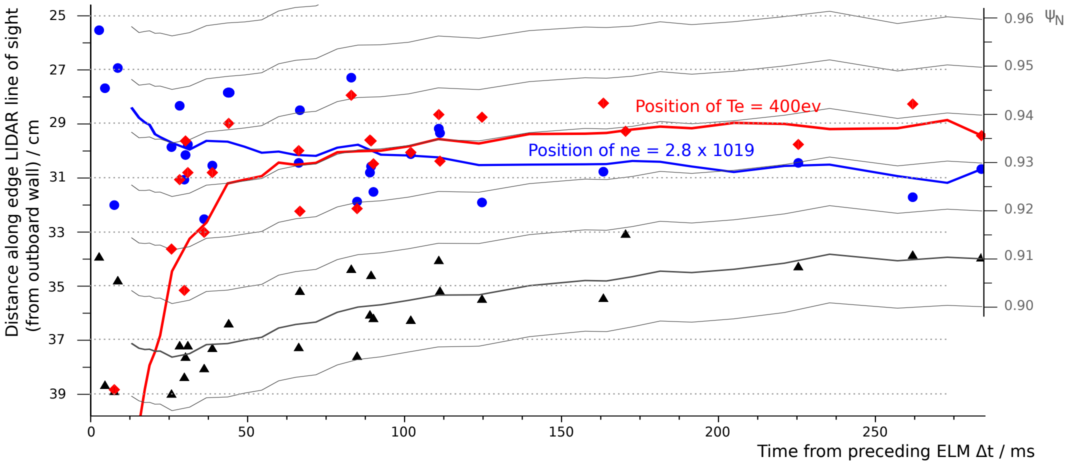

Figure 5.22.: Position in real space along edge LIDAR line of sight of $\psi_N = 0.91$ (black triangles), $n_e = 3.0 \times 10^{19} m^{-3} \approx \frac{1}{2} n_{e,ped}$ (red diamonds) and $T_e = 400eV \approx \frac{1}{2} T_{e,ped}$ (blue circles). 10 point moving averages are shown for each, and for intervals of 0.1 in ψ_N.

5.8.4. Shapes and marking

It is clear from figure 5.20 that most of the density profiles can be represented by the modified tanh profile discussed earlier. The T_e profiles however, can not be represented by this form. Early in the ELM cycle, they can be very flat, showing no clear pedestal beyond the noise. Late in the ELM cycle (panels 5 and 6 of figure 5.21), the T_e profile has a slow transition from the gradient at mid-minor radius ($\psi_N \approx 0.8$) into a steep, constant gradient in the last 1-2 cm. This pedestal gradient is maintained down to less than $100eV$ and beyond the minimum T_e at which edge LIDAR can observe TS light. Inversions using both the mtanh and the linear pedestal models were attempted for these pulses but the edge LIDAR data at the top of the pedestal is sufficiently strong that it dominates the improperly fitting model, giving a much wider pedestal than is seen in the free-form profile inversions. In general, these fits were too strongly effected by noise and stray light.

To quickly quantify the pedestals without a well fitting model, the top and foot of the pedestal in each of the free-form profiles was marked by hand. To reduce unintentional biasing, this was done in an entirely random order, with both the x and y axes obscured (no scales). In many cases, the profiles did not exhibit a clear enough pedestal shape to mark, so the range in ψ_N which might be considered the top of the pedestal was marked instead of a fixed point. Figure 5.23 shows the marks for an n_e profile and for both clear (mid-cycle) and unclear (late) T_e profiles.

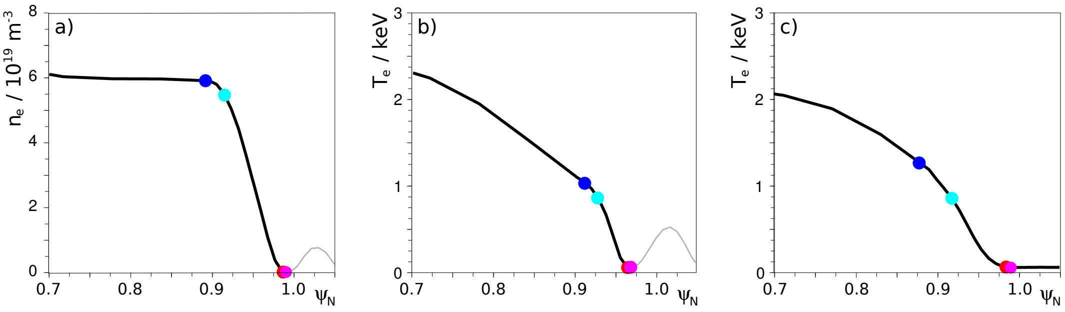

Figure 5.23.: Marking of range of possible locations of the pedestal knee (blue) and foot (red) in profiles for a) a typical n_e pedestal b) a relatively clear T_e pedestal and c) a unclear and difficult to characterise T_e pedestal.

The heights, widths and gradients from all 28 time points are shown in figure 5.24, plotted against (and coloured by) time after the preceding ELM and with error bars calculated from the marked ranges. Linear fits are shown for $\Delta t > 50ms$ for all plots and for $\Delta t < 50ms$ for plots involving T_e (right).

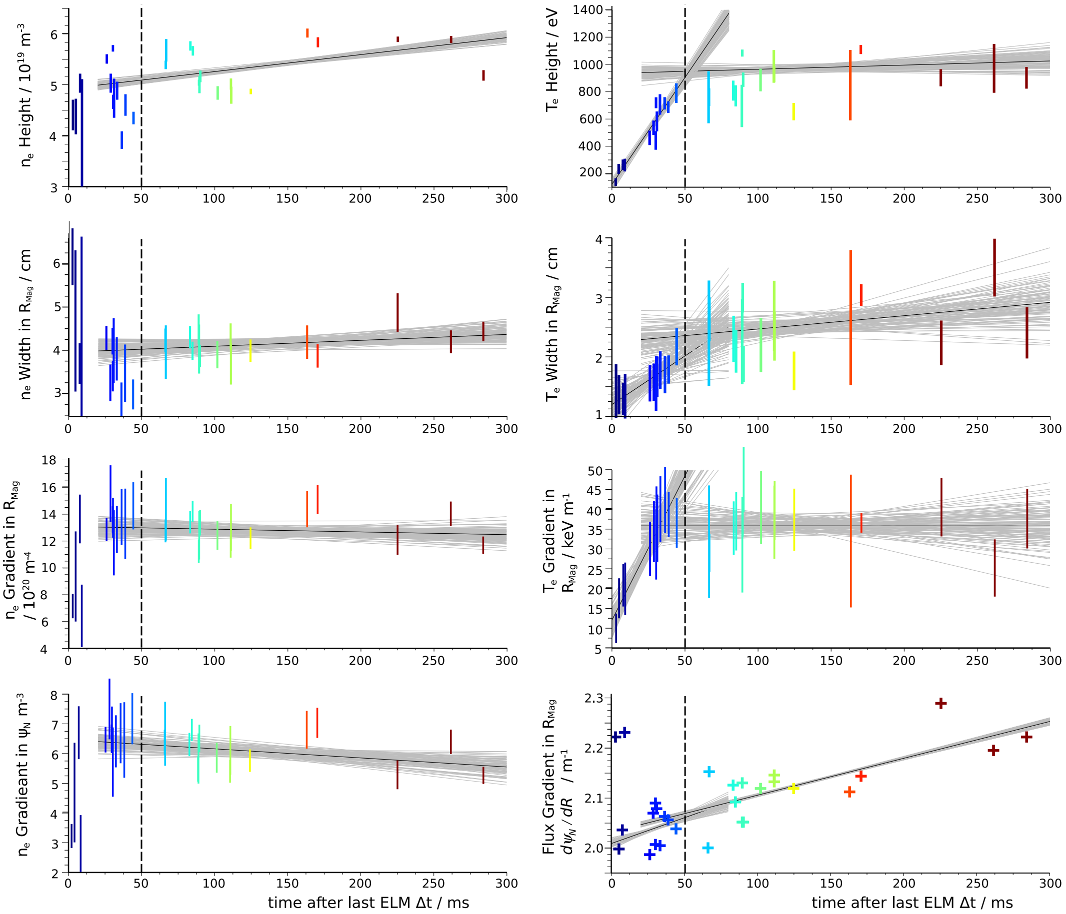

Figure 5.24.: Evolution of n_e (left) and T_e (right) pedestal parameters over ELM cycle from marking of 28 edge LIDAR time points in 6 similar pulses (see table 5.3). 1st row: Pedestal height, 2nd row: Pedestal width in magnetic mid-plane real space (R_{Mag}), 3rd row: Pedestal gradient (R_{Mag}). The last row gives the n_e pedestal gradient in ψ_N space, in which it was marked, and the mapping gradients $d\psi_N/dR$ (from EFITJ). Points are plotted against and coloured by time from preecding ELM (Δt). Linear fits (black) and uncertainty (grey, as samples) are shown for $\Delta t > 50ms$ for all plots and for $\Delta t < 50ms$ for plots involving T_e.

5.8.5. Width and gradient evolution

The difficulty in defining the pedestal leads to an uncertainty that is quite large, especially for late cycle T_e width and gradient, but some trends can still be seen. The variation beyond those trends is greater than the uncertainty, which suggests there are some significant differences between the plasma at the same stage in different ELM cycles, as these profiles came from different cycles and different pulses.

Despite these considerations, it is possible to see some of the features discussed earlier and a few more beyond those. The T_e pedestal height rises rapidly and reaches close to its late ELM cycle values after only $40ms$ while the density pedestal reaches $4 - 6 \times 10^{19} m^{-3}$ almost immediately. There is a subtle trend of slowly increasing density and temperature after $50ms$ and although this could easily be an artifact of the cycle and pulse variations, it is supported by the high temporal resolution ECE and interferometry.

The early $(20 - 50ms)$ data, seems to indicate that the T_e gradient returns to its pre-ELM value very rapidly, and that the width and height increase together, rapidly before $50ms$ and slowly, if at all, after that. The density pedestals, being much clearer for most of the cycle, have a smaller uncertainty for all but the very early $(\Delta t < 15ms)$ time points. Again, a possible slow increase over the rest of the cycle might be present, but with the gradient more or less constant. Most of figure 5.24 shows the R_{mag} variation of the profiles which can be different to that in normalised flux as the plasma moves through the ELM cycle. The effect is not as strong as discussed earlier for the position, since here it is only effected by differences between the flux compression at the outboard midplane and on the line of sight. The outboard midplane mapping gradient $d\psi_N/dR$ is plotted in the bottom-right graph and shows a gradual increase throughout the ELM cycle, as the plasma β increases, increasing the Shafranov shift and compressing the outboard flux surfaces. However, this is a change of only around 20% and so the variation of width and gradient is not clearly different in ψ_N beyond the uncertainty. However, from the linear fit, there is a suggestion that $dn_e/d\psi_N$ falls as $d\psi_N/dR$ increases, retaining the same dn_e/d_R.

5.8.6. Conclusions and scope for future investigation.

This investigation was limited in scope and performed on a small number of pulses (6) for a single plasma configuration. The trends seen are on the very limits of the diagnostic resolution and are comparable with the variation of the plasma between ELMs, so the conclusions drawn are not presented as proven results. For this, the work must be extended to consider a larger data set. However, this study does serve to demonstrate the capability of the inversion to extract very accurate individual pedestal profiles from the edge LIDAR system, which is not possible with the ELM-cycle averaging of profiles performed elsewhere. The edge LIDAR data suggests that the modified tanh profile is not a perfect model for the T_e pedestal. With more high signal to noise shots like those analysed here, it might be possible to construct a better ad-hoc parameterisation, which could then be used with the remainder of the data (over 16000 time points with 7800 in type-I ELMy H-mode), to study the pedestal width, height and gradient variations, both within the ELM cycle and between different plasmas. This work has also indicated that two seperate aspects of the density pedestal (position and gradient) might have a clearer behaviour in real space, than in the magnetic geometry. A similar effect has been seen on MAST [79] where, in a double-null configuration (upper and lower X-point), the inboard and outboard density profiles have a consistent shape in R_{mag}, but the temperature profiles have a consistent shape in ψ_N. This does also indicate that the assumption of constant n_e on flux surfaces may not be valid. This should be investigated, and while it would be relatively trivial to exchange the plasma model to include variations within a surface, how much could be inferred with such flexibility is not immediately clear. As was shown in section 5.7, the effect of the flux surfaces on the shape (and hence width and gradient etc) inferred is quite strong, so the inversions here should also be performed with the full Current Tomography model, allowing the flux surfaces uncertainties to also be included.

6. Bayesian Equilibrium

6.1. The Equilibrium Prior

6.1.1. Current Tomography versus equilibrium solutions.

It was explained in section 1.3 that external magnetic coils alone provide insufficient boundary conditions to solve the equilibrium equation without making assumptions about the form of the current and pressure distributions. It was shown experimentally in section 4 that the solution found this way can be sufficiently inaccurate to complicate and confuse the analysis of diagnostics. Alternatively, combining those diagnostics with the magnetics model and performing a single inversion (Current Tomography) provides a rigorous way to allow for and reduce flux surface uncertainty without assuming equilibrium but, as was shown in section 4.2.3, the uncertainty found this way, without adding further diagnostics, is often larger than the error in the equilibrium solution. The choice is between an inaccurate estimate with an uncertainty which includes the reality and a more accurate (but still critically incorrect) answer with no uncertainty.

Since we do expect that the real plasma is at least close to equilibrium, what is really desired is the PDF of all plasmas which are consistent with the data and are close to equilibrium solutions. The fact that the equilibrium cannot be solved outright with only magnetic diagnostic data does indicate that the profile smoothing priors etc, that are chosen will have a great effect but, unlike the fixed solution, adding more diagnostic data will always override these assumptions if the priors are sufficiently weak. Previously, a PDF of a range of possible equilibria consistent with various measurements has been found using Bayesian methods for a Stellerator [37]. This was done by interpolating a set of equilibria previously calculated by a traditional solver, restricting the PDF to plasmas which follow the assumptions made by that solver.

The objective here, is a PDF of every possible plasma which will also describe what really can be inferred about the plasma equilibrium from just magnetic diagnostics, which is a strongly debated topic.

6.1.2. The simple Grad-Shafranov difference prior

Including the assumption of equilibrium in the full CT model is conceptually simple. A parametrisation for the *equilibrium pressure* $p(\psi_N)$ and poloidal current flux $f(\psi_N)$ is added to the model and the difference δ^{GS} between the toroidal current density j_ϕ and the right hand side of the isotropic pressure flow-free Grad-Shafranov (GS) equation 1.15 calculated at a series of points throughout the (R, Z) plane. A prior PDF $P(j_\phi,\, p,\, f)$ over this difference is included in the posterior. It can have arbitrary shape but must be centred at 0 difference and is usually chosen as a simple Gaussian:

$$\delta^{GS}(R,Z) \;=\; j_\phi(R,Z) \; - Rp'(\psi_N) \; - \frac{\mu_0}{R} f(\psi_N) f'(\psi_N) \qquad [\,\psi_N = \psi_N(R,Z)\,] \tag{6.1}$$

$$P\left(j_\phi, p', ff'\right) \;=\; \prod_{i=0}^{n}\prod_{j=0}^{m} \mathcal{G}\left(\delta^{GS}(R_i,Z_j);\; 0,\, \sigma_{equi}\right) \tag{6.2}$$

The equilibrium pressure function p may be related n_e, T_e, T_i etc via other prior assumptions at a later stage.

If the likelihood function of only the magnetic diagnostics $P(D_m|j_\phi)$ is included, the posterior PDF $P(j_\phi,\ p',\ ff' \mid D_m)$ will now include all plasmas consistent with the magnetic coils that are equilibrium solutions or are close to one. The uncertainty represented by this posterior includes the uncertainty due to the uncertain magnetic measurements but also that due to the degeneracy in the GS equation.

6.1.3. Beam average and net force priors

As with all prior PDFs (and the corresponding assumptions in frequentist methods), there are some arbitrary choices in this specification which introduce complications. The parametrisation of j_ϕ by uniform beams is actually incompatible with equilibrium since the beams will each relate to small overlapping intervals of p and ff' in ψ_N space. For the GS equation to be exact, p' and ff' must be uniform over each interval and because they overlap, must be uniform everywhere. The prior assumption of small GS difference over multiple points on each beam inherently implies prior belief that p' and ff' are flat. Many parametrisations (e.g. linear or cubic 2D interpolations, delaunay triangulation interpolation etc) were tried but each has similar issues. The simplest (and computationally least intensive) solution is to require only that the average of $Rp' + (\mu_0/R)ff'$ over each beam must be close to j_ϕ:

$$\delta_i^b \;=\; j_\phi(R_i,Z_i) \; - \frac{1}{A_i}\iint_{A_i} Rp'(\psi_N) \; + \frac{\mu_0}{R} f(\psi_N) f'(\psi_N) \; dRdZ \tag{6.3}$$

$$P\left(j_\phi, p', ff'\right) \;=\; \prod_{i=0}^{N.beams} \mathcal{G}\left(\delta_i^b;\; 0,\, \sigma_{equi}\right) \qquad \left[\, A_i = i^{th}\text{beam area} \,\right] \tag{6.4}$$

This constraint is similar to requiring that the net force on each beam is small - without assumption about compression or expansion forces inside each beam. It differs from this only by the factor $\nabla\psi/R$ that is present in equation 1.14. This term gives a different relative weighting between different regions of the plasma, a distinction which is another arbitrary choice of the prior assumption.

6.1.4. Scrape off layer and private regions

The situation outside the LCFS is more complex as each surface is connected to the first wall and it can not be assumed that areas of the same ψ_N are connected. However, to maintain a practical parametrisation this is ignored and both p' and ff' remain functions of ψ_N for $\psi_N > 1$. For private regions (usually between X-point and divertor) which have $\psi_N < 1$ but are outside the LCFS, the flux is reflected through the Separatrix $\psi_N \longrightarrow (2 - \psi_N)$ to require that they have similar current to the SOL. The assumption that both p and ff' are small for $\psi_N >\ 1.05$ is then added as a Gaussian prior PDF. This allows current at the LCFS, Separatrix and a small way into the SOL but discourages large currents far out from the plasma edge, which are not expected. This in turn requires j_ϕ to be small in these areas through the GS prior. These assumptions are based on general observations of the SOL currents at JET[80] and on other tokamaks[81] which show $j_\|$ of at most 1% of the bulk plasma current. Measurements of electron and ion temperature and density do not generally show large pressure gradients in the SOL which directly supports the prior assumption of small p' (and hence small $j_\perp$).

6.1.5. Full Posterior

The full posterior, which is explored in the remaining sections of this chapter is:

$$P(j_\phi, p', ff' \mid \underline{\mathrm{D}}^{mags}) = P(\underline{\mathrm{D}}^{mags} \mid j_\phi) \cdot P\left(j_\phi, p, f\right) \cdot P(p') \cdot P(ff') \cdot P(j_\phi) \qquad (6.5)$$

The likelihood distribution $P(\underline{\mathrm{D}}^{mags} \mid j_\phi)$ is the multivariate Gaussian of the magnetic diagnostics data around the prediction from the Current Tomography model[38]. The second term is the beam averaged equilibrium prior of equation 6.4, with $\sigma_{equi} = 50kAm^{-2}$, chosen as $\sim 1\%$ of the typical average plasma current density. The profile priors, $P(p')$ and $P(ff')$ can be used to apply weak regularisation, if desired, and are discussed separately in each section. Finally, the beam current density prior $P(j_\phi)$ can now be assigned a uniform prior. The CAR prior used in the earlier current tomography based work is now not required, as the combination of the equilibrium and profile priors are sufficiently constraining to make the posterior finite, although still very broad in some directions.

6.2. Simulations I - Simple near-circular limited plasmas.

To demonstrate the method where an approximate equilibrium solution is known to exist, and to investigate the problem without the presence of noise of the magnetic measurements, this section gives posteriors inverted from simulated magnetic data. It also numerically confirms some of the known theoretical results regarding what can be recovered. The simulated data is based loosely on a real JET pulse and time point to maintain a realistic plasma shape, total current and total pressure.

6.2.1. Circular low-beta EFITJ reconstructions

In this case, j_ϕ is taken from the standard EFITJ solution for a limited L-mode low-beta plasma with relatively circular shape. The forward model is used to predict noiseless magnetic measurements and an inversion is then performed on the simulated data to attempt to recover the original j_ϕ, p' and ff'. Figure 6.1 shows the original profiles and profile marginals of the posterior and figure 6.2a shows the associated flux surfaces. For both the simulation and reconstruction p' and ff' are assigned profiles of 6 knots fixed at values of ψ_N that result in roughly equal spacing in $R(Z = Z_{mag})$ and j_ϕ is parameterised with a 30x30 set of current beams. The normal (realistic) uncertainties are used for the magnetic measurement likelihood distribution and the profile and current density priors, $P(p')$, $P(ff')$ and $P(j_\phi)$ are all assigned uniform priors. The posterior was explored using the LGI bootstrapped adaptive proposal Metropolis Hastings method covered in section 2.4.5.

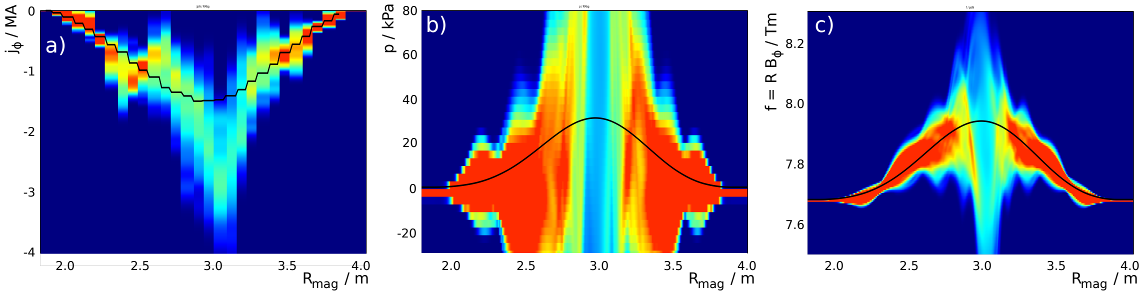

Figure 6.1.: Posterior PDF inferring equilibrium from simulated signals for JET magnetic diagnostics. a) toroidal current j_ϕ, b) pressure p, c) poloidal current flux f

The clearest result is that the whole equilibrium is more uncertain towards the centre of the plasma which is an expected result since the plasma is almost circular. In the limit of infinite aspect ratio (a completely cylindrical plasma) the external magnetics sensors alone would be unable to distinguish between current rings at different radii. The second striking result is that p is very uncertain, especially compared to f. This can be understood by recalling that only poloidal magnetic measurements are used here and they provide direct information only about j_ϕ. p' and ff' are given information only by their coupling to j_ϕ via the GS equation. In this low-beta plasma, the magnitude of the Rp' term in GS equation is much less than that of j_ϕ and ff'/R. A small fractional uncertainty in j_ϕ implies an uncertainty on the two terms which is a large fraction of p' but a small fraction of ff'.

Although it cannot be clearly seen in the profile marginal plots, the large uncertainties come principally from the possibility of oscillations in various quantities which have a small effect on the magnetic measurements and therefore cannot be excluded as possible solutions. To demonstrate this, figure 6.2 compares the flux surface uncertainties from this inversion with another using measurement uncertainty 100 times smaller. The edge surfaces show much smaller uncertainty in the latter case, showing that it comes mainly from the measurement uncertainty. The uncertainty in central flux surfaces comes instead from the problem degeneracy and so is not effected by the change in measurement uncertainty.

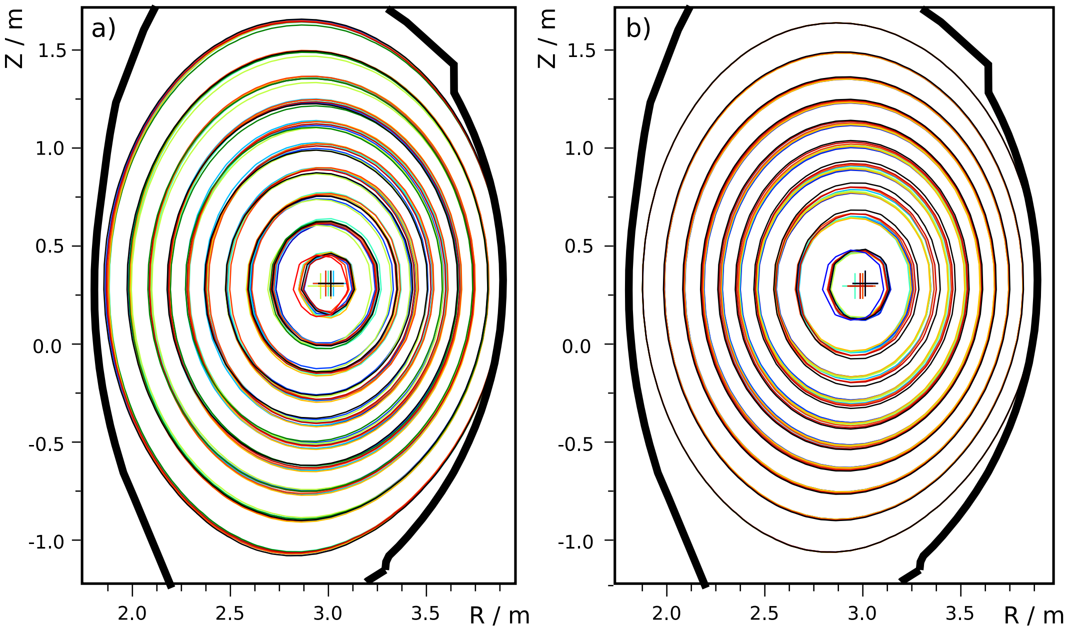

Figure 6.2.: Contours of normalised flux in several posterior samples from inversion of simulated magnetics data with a) full and b) 1% of normal uncertainties.

6.2.2. Determination of moments β_θ, μ and l_I

As outlined in section 1.3, there are quantities which can be related directly to loop integrals of the poloidal magnetic field around the plasma boundary. Two of these are the Shafranov integrals which relate to the useful quantities β_θ, μ and l_I as $s_1 \sim 3\beta_\theta - \mu + l_I$ and $s_2 \sim \beta_\theta + \mu + l_I$ [15]. Since the loop integrals relate almost directly to the edge measurements, the uncertainties on s_1 and s_2 are tied closely to the measurement uncertainty. If these are accurate, all possible equilibria within any remaining uncertainty should follow the relationships:

$$\mu \sim \beta_\theta + c_1, \qquad l_I \sim -2\beta_\theta + c_2, \qquad l_I \sim -2\mu + c_3$$

What else can be be known about β_θ, μ and l_I within this, depends on the higher order moments of the magnetic information and so in general will be less accurate then the above relationship, especially for more circular plasmas.

Figure 6.3 shows the correlation between the three quantities in the simulation posteriors using the full (blue) and 1% (red) magnetic measurement uncertainties. In both cases, it is clear that greater part of the uncertainty is through this partial degeneracy, as the scatter is greater along the directions given by the above relationships (green). However it shows that β_θ, μ and l_I can still be determined individually, albeit with large uncertainties.

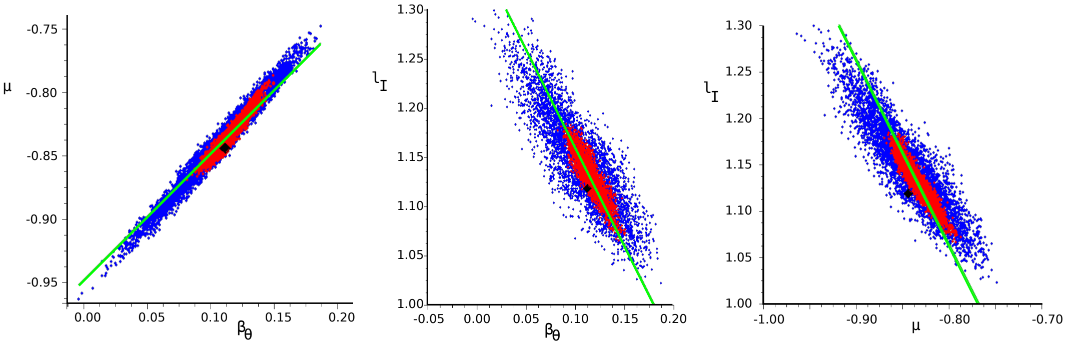

Figure 6.3.: Normalised moments β_θ, μ and l_I for samples of the posterior using normal (blue) and 100 times reduced (red) measurement uncertainty. Black diamonds show values for the original simulated equilibrium and the green lines show the gradients of the expected theoretical degeneracy for the entirely cylindrical case.

The recovery of this theoretically anticipated partial degeneracy shows that posterior exploration is being completed correctly and demonstrates the power of this method to represent the full range of equilibria consistent with the magnetic measurements.

6.3. Simulations II - High resolution reconstruction.

What can be inferred about j_ϕ, p' and ff' beyond the global low-order moments is strongly debated. Analytical calculations show little can be inferred for cylindrical and possibly near circular cases while numerical work often shows that at least a few higher order moments can be found accurately in elongated plasmas. Some of the previous literature [16] effectively dismisses the numerical work, stating that the p' and ff' profiles can never be inferred beyond integrals of them. There is some conceptual confusion here since, while it is absolutely true that only integrals can ever be known, it is not a speciality of the GS equation or the magnetic measurements. It is never possible to completely determine spatial profiles from a finite number of real measurements when the measurements are integral in nature, which is almost always true. There is no discrete difference between the inference of electron temperature from Thomson scattering spectrometers that collect light over 5mm and one from magnetic measurements integrating current over the whole plasma. There are no 'point measurements' in reality and so it is always possible to find many perturbations, usually oscillation like, which cannot be detected. Without the prior assumption that these are not present, the answer to the question of what can be inferred about the quantity at a single point is simple - nothing.

In summary, exactly what can be inferred from magnetic measurements remains unproven and the Bayesian methods provide a way to, in principle, extract and represent this for each case. Profiles are used for a parametrisation simply as a basis on which to work and it is what is assumed and what can be inferred about them, the PDFs, that is of interest.

To investigate, a simulated plasma was created from the EFITJ solution for an existing JET diverted H-Mode plasma with some arbitrary local deformations forced into p'. The nearest equilibrium is found and the magnetic coil predictions taken as simulated data D_m. To represent the local features, 15-knot profiles are used for p' and ff' and 4.5cm width current beams used for j_ϕ.

The simulated data are inverted to find $P(j_\phi,\ p', ff' \mid D_m)$ using 30-knot p'/ff' profiles to match the real situation in which the scale lengths of the features are not known. j_ϕ retains the 1510 4.5cm beams as in the original profile. Initially, the profiles priors $P(p')$ and $P(ff')$ are assigned weak first differential smoothing with $\sigma_{d(p')/d\psi_N} = 10^{10} Am^{-3}$ and $\sigma_{d(ff')/d\psi_N} = 10^{10} \mu_0^{-1} Am^{-1}$

6.3.1. The Iterative Linear Scheme

Unfortunately, the 1510 current beams required to adequately describe the complex geometry of j_ϕ causes the posterior dimensionality to be so high that the full non-linear optimisation and exploration methods are extremely time consuming and difficult to use. Even running in parallel over 100 modern machines, the Genetic Algorithm can take up to a few days to reach a satisfactory MAP estimate and it is the fastest of the stable non-linear algorithms tried so far.

With the pure current tomography, the priors were entirely linear and this problem was avoided by using the LGI. The addition of the GS prior breaks the linearity. However, if the functions p' and ff' are used directly and parametrised by a simple interpolation of knots, then δ is linear in those parameters at fixed j_ϕ. j_ϕ itself appears linearly in the first term but also enters δ through ψ_N introducing a non-linearity which is only strong if p' and ff' have large gradients. In practice, the posterior is locally well enough approximated as linear, that the LGI gives a good estimate of the direction of higher posterior probability. The following iterative procedure uses this to gain a good estimate of the global MAP after only a few minutes:

1. Obtain j_ϕ from standard Current Tomography LGI with CAR prior or from the standard EFITJ solution).
2. Remove the CAR prior and add the GS prior.
3. Calculate linear coefficients at present position and find estimated MAP using the LGI.
4. Search in the direction of the new estimate from the current position for the highest true posterior.
5. Move to the position with the highest real posterior.
6. Repeat from 3.

It is necessary to assign a Gaussian for the beam current density prior $P(j_\phi)$, in order for the LGI to be used. An uncorrelated Gaussian, centred at 0 with $\sigma_{j_\phi} = 10^{10} Am^{-2}$ is used so that typically, $\sigma_{j_\phi} > 10^3 j_\phi$. This is sufficiently weak that it does not strongly effect the posterior.

After many iterations, the approximate MAP will remain stationary and a (at least local) maximum in the true posterior has been found. The procedure is mathematically similar to that employed by many equilibrium solvers but because each stage only moves to higher values of the true posterior, the procedure can not diverge. The result is always a higher posterior probability, close to equilibrium and matching the data D_m well, even if it is not physically sensible in other ways.

While this method eventually converges on the MAP, the covariance obtained by the LGI does not give a good estimation of the posterior shape and size. Investigation of the uncertainty should be carried out by non-linear sampling methods but again, the high dimensionality makes such exploration extremely difficult. Much time and effort was expended in the investigation and development of the Metropolis Hastings, nested sampling[41] and a few other exploration methods but a practically usable algorithm was never found. The best results were achieved using 100 parallel Metropolis Hastings chains, each individually scaling a proposal covariance periodically derived from all the global samples but even this only began to explore solutions far from the initial MAP after several days. The exploration was incomplete so the results are not presented here but the work does suggest that representative samples could be achieved within a workable length of time with further development of the algorithm and/or more computational power, giving full rigorous determination of all possible high-resolution equilibria from experimental data.

6.3.2. Exploration by prior variation

Abandoning rigour temporarily, it is possible to gain some insight into what profiles are possible by using the iterative LGI procedure with a range of priors on p' and ff'. If these profile priors are kept relatively weak, the MAP estimates are always solutions close to both equilibrium and the data. The profile priors only fill in the information not provided with confidence by the data and so varying them gives an idea of the uncertainty present, albeit a qualitative one.

Figure 6.4 shows the original profiles and the iterative LGI MAP estimates using first or second differential smoothing priors (equations 4.1 and 4.2) with a range of strengths on the knots of p' and ff'. A smoothing prior is not required for j_ϕ, since it is indirectly regularised by those on p' and ff' via the equilibrium prior.

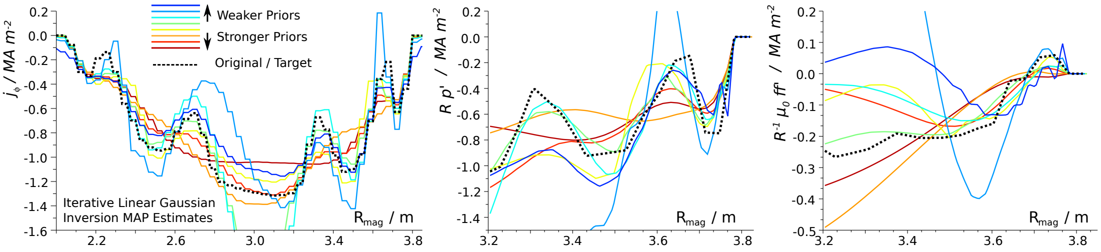

Figure 6.4.: Original generated (black) and MAP estimates of j_ϕ, p' and ff' profiles, from the iterative linear scheme of section 6.3.1 with different smoothing priors on p' and ff' (coloured roughly by relative strength).

The reconstruction is often qualitatively reasonable, showing always the general shape and in many cases the local features but it is clear the profile priors do have a strong effect. The weaker priors allow large amplitude oscillations while strong priors flatten some of the features. Despite this variation, the magnetics predictions never stray outside 0.1σ of the simulated data and δ^b is always close to 0 showing good proximity to equilibrium. The unsurprising conclusion is that even fairly accurate magnetic data is insufficient to completely reconstruct every detail of the original profiles.

The variation in j_ϕ at the plasma core is much larger than at the edge implying the edge is better diagnosed. This matches that seen in the properly explored uncertainties of section 6.2. In contrast to those results, p' appears better diagnosed than ff' though this is consistent since the plasma is much less ferromagnetic here. The $\mu_0 ff'/R$ term is much smaller in magnitude than Rp' so a large local uncertainty in j_ϕ gives a larger uncertainty in ff' than in p'.

One feature that is always present is the initial rise in the profiles at the plasma edge. The magnitude of j_ϕ at $R = 3.7m$ and $R = 2.1m$ varies by less than 20% around its target value and although this relates to a $\sim 40\%$ local uncertainty in p' and even larger still in ff', the feature is always present in all three profiles indicating that the data holds enough information to be certain of its presence and relatively confident of its amplitude in j_ϕ and p'.

6.3.3. Inference of Pedestal Current

While rigorous uncertainties must await further development of the sampling methods (or more powerful computers), these result are encouraging. The amplitude of the current at the edge of the plasma is of great interest in the study of the H-Mode pedestal as the pressure gradient ∇p and/or field parallel current $j_\parallel$ are thought to play a role in the triggering of ELMs[82]. Of the ion and electron components of p, p_e is reasonably well diagnosed (e.g. from chapter 5) but the ion pressures are harder to determine. Bulk and impurity ion temperatures T_i can be measured by the Charge Exchange Recombination Spectroscopy (CXRS) diagnostic, which observes emission due to interaction of those ions with the neutral beam particles. The ion and impurity densities are the more difficult part and typically involve measurement of the effective charge Z_{eff} and assumptions of the impurity species present. Measurement of $j_\parallel$ has attracted a lot of experimental effort recently including measures of Motional Stark Effect (MSE) [83], also from the main heating neutral beams, or the Zeeman Effect on independently injected Lithium[84]. Such measurements often suffer poor signal to noise, insufficient temporal resolution to observe in-between ELMs or sensitivity to other unknowns like the electric field. A new method, with very high resolution, is currently being developed at MAST [85]. It should be noted that none of these measure $\underline{j}$ directly but instead measure the magnetic field components or the ratio B_θ/B_ϕ (i.e the pitch angle) and use an equilibrium solver and/or simple assumptions to calculate the current. In every case, the determination of the $\underline{j}$ is a complex inference problem.

Given the results of figure 6.4 it is unlikely that the magnetics data alone will ever be used to examine the fine structure of the pedestal current, or even whether or not other structures exist further into the plasma but it may be possible to infer the parallel or perpendicular currents or something about the relationship between them, especially if prior information about the shape is included. As a simple investigation, a series of equilibria were generated that were similar to figure 6.4 but with 5 different values of p' and 5 of ff' in the edge peak. From the predicted magnetics data of each, the LGI MAP estimate was found using a simple parametrisation for p' and ff' of only 3 knots and a small Gaussian peak at the edge of each profile. Given the strong parameterisation, the priors $P(p')$ and $P(ff')$ can be uniform, with effectively infinite range. Figure 6.5 shows the original and the MAP estimate profiles for 4 of the 25 generated equilibria.

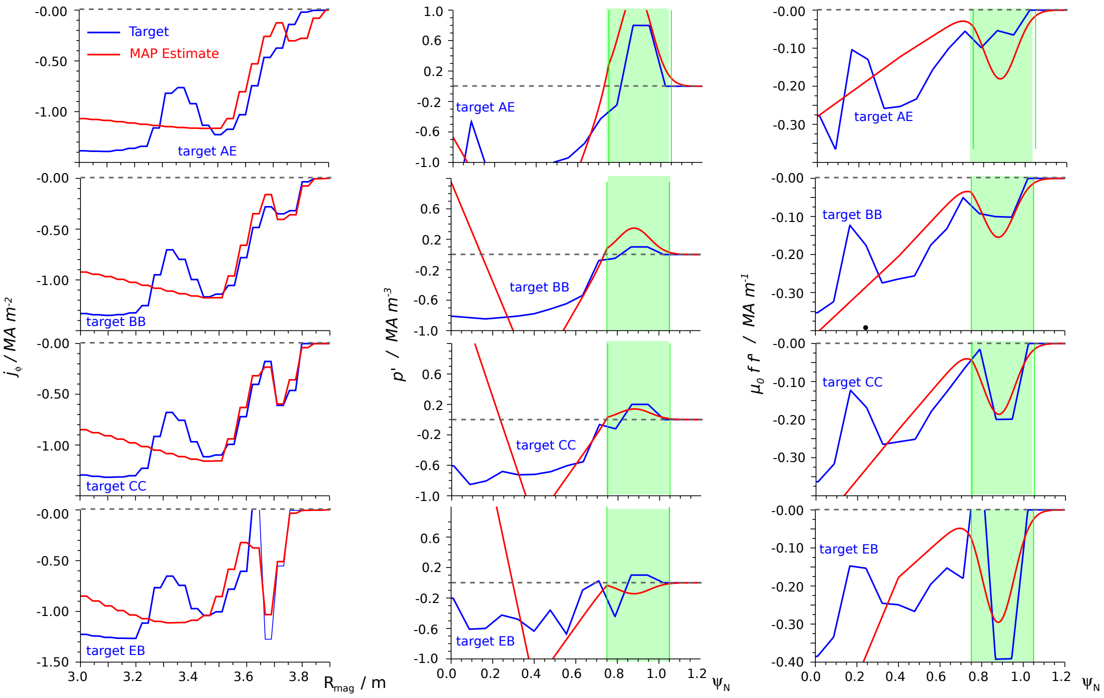

Figure 6.5.: Original (blue) j_ϕ, p' and ff' profiles for a selection of edge currents and iterative LGI MAP estimates (red) using a simple 3 knot line and Gaussian edge peak model.

As with the prior variation, j_ϕ at the edge is recovered well in each case while the accuracy of the p' and ff' MAP estimate varies. In some cases, the fact that the prescribed Gaussian peak can not match the target profile is evident and this reflects the real situation since the fine detail of the Tokamak edge current is still not well known. It is likely however that integrals over the edge region should match well. Since it is the field parallel and perpendicular components that are usually discussed in ELM theory, the flux surface average of these, integrated over $0.75 < \psi_N < 1.00$ (shown in green in figure 6.5) are calculated for each target and each MAP estimate. Figure 6.6 shows this edge current for all 25 targets and estimates.

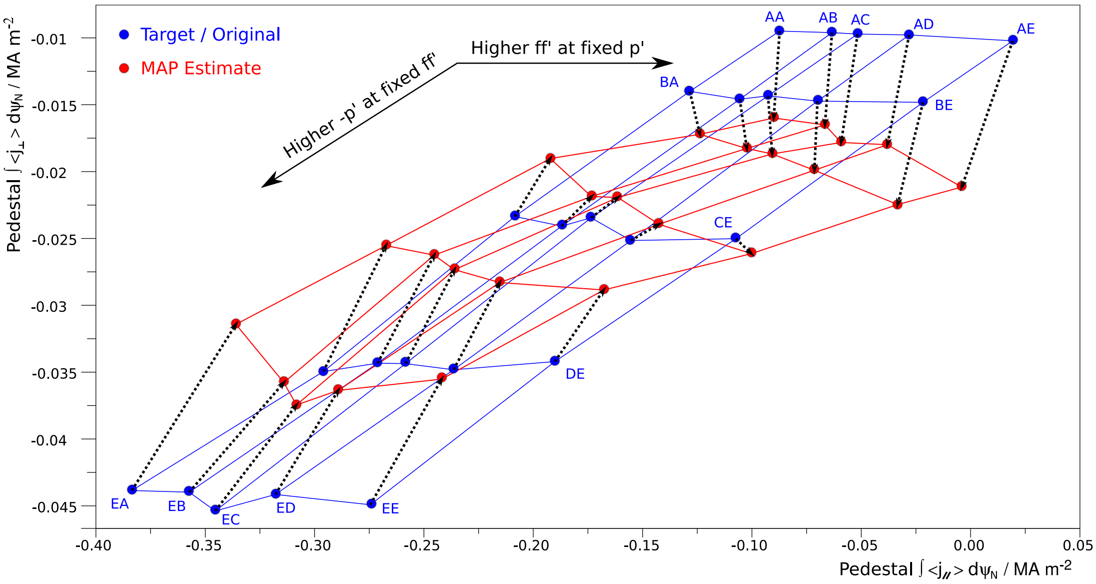

Figure 6.6.: Flux-surface average parallel and perpendicular currents $< j_\perp >$ vs $< j_\parallel >$, integrated over the pedestal region $(0.75 < \psi_N < 1.0)$ for 25 target simulated equilibria (blue) and iterative LGI MAP estimates (red).

The MAP estimates are in some cases a relatively large distance from their target which indicates that the method will not immediately provide an accurate measurement of the magnitude of both components. However, the fact the 2d spread is preserved suggests both quantities might be inferred to some extent. If there were one degree of complete freedom (not necessarily along $< j_\parallel >$ or $< j_\perp >$) on which the magnetic data and equilibrium provided no information, the required information would be filled by the prior and since the priors were fixed for all 25 points, the results would follow a line and lose their 2d nature. For a rigorous account, this must of course be re-examined when the full posterior can be explored properly but this investigation suggests it may be possible to obtain at least the qualitative variation of the plasma edge in the $(< j_\parallel >, \nabla p)$ and since only magnetic data is used, this could be performed at their full temporal resolution, which on JET is much faster than the typical type-I inter-ELM build up time. At the very least, the ∇p variation can be compared with what is known about p_e and p_i from the kinetic diagnostics.

6.4. Reconstructed H-Mode Equilibria

6.4.1. LGI MAP Results and prior selection

After rejecting suspected failed magnetic diagnostic coils (see Appendix B), the iterative LGI is applied to find the MAP estimate for real data from a typical JET type-I ELMy H-Mode pulse. As with the reconstruction of the simulated profiles, the exact choice of prior effects the MAP, regardless of how weak it is, in directions in which the data provides no information and hence the posterior is very flat. Figure 6.7 shows the MAP estimates under a verity of smoothing priors.

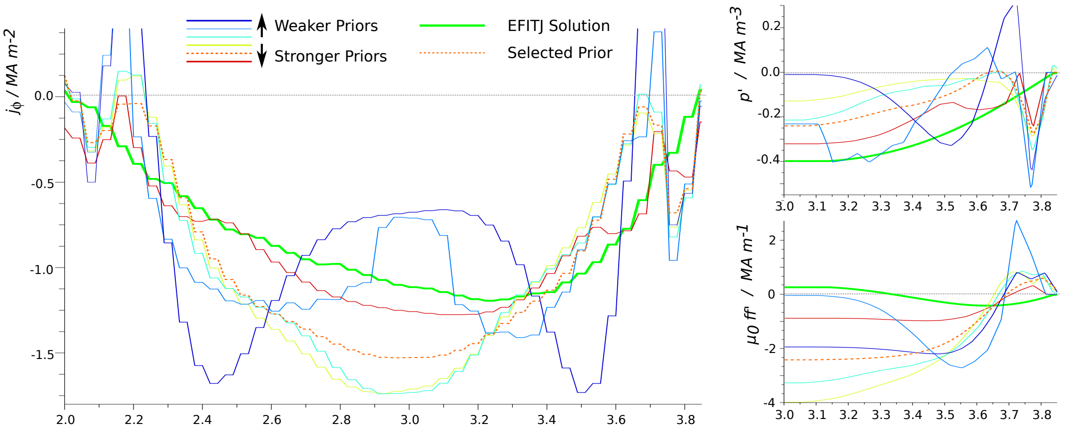

Figure 6.7.: MAP estimates of $P(j_\phi, p', ff'|D_m, \text{Equilibrium})$ from the iterative linear scheme of section 6.3.1 with different smoothing priors on p' and ff' (coloured roughly by relative strength). Magnetics data D_m taken from pulse 78601 - a type-I ELMy H-mode at 58.530s, mid-way between two ELMs.

While the plasma core shows heavy dependence on the prior, the peak in j_ϕ associated with the H-Mode pedestal is seen clearly in every case, despite the tendency of smoothing priors to discourage such features. There is a variation in the amplitude of this peak in all three variables but the profiles with large peak edge current also show a positive region of p' (i.e. a non-monotonic pressure profile) and/or a large reduction in j_ϕ in the plasma core. While such profiles, known as *current holes*, have been observed in some special cases[86], it is not expected in this pulse where only the Ohmic and neutral beam systems are providing heating/current drive. Non-monotonic pressure profiles are not observed by any kinetic diagnostic and are not expected to occur. Excluding the priors which result in these unexpected features leaves four which all give approximately the same amplitude of edge current. Figure 6.8 shows j_ϕ, flux surfaces, q profile (Safety factor) and pressure profile for one of them, selected arbitrarily, compared to the standard EFITJ solution. The selected prior in is shown as broken orange in figure 6.7.

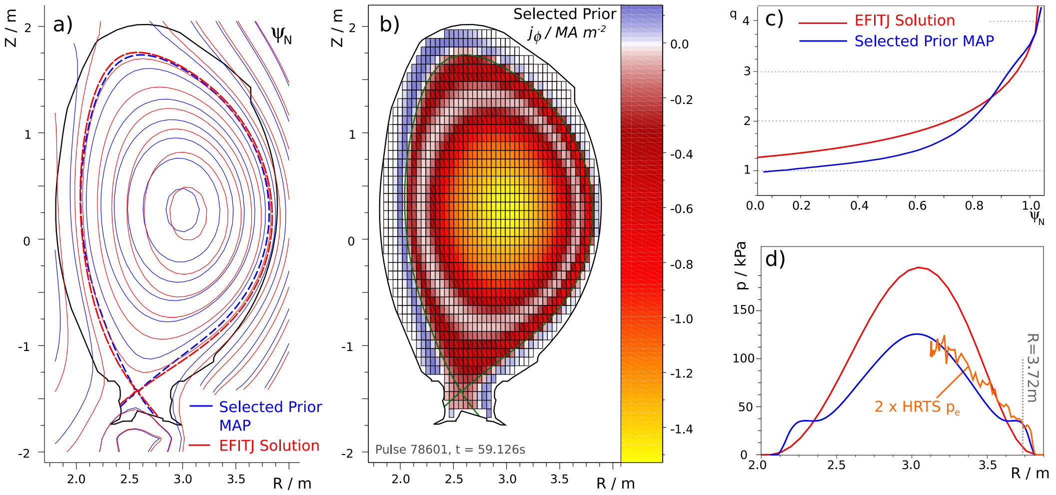

Figure 6.8.: a) Flux surfaces, c) q profile (Safety factor) and d) pressure profile for selected prior compared to standard EFITJ solution. b) Full 2D $j_\phi(R, Z)$ for selected prior. The electron pressure profile is also shown doubled for comparison (i.e assuming $T_i = T_e$ and $Z_{eff} = 1$).

Where it can be assumed that $p_i = p_e$, the MAP estimate p profile should match $2p_e$ calculated from electron kinetic measurements, so the Thomson scattering standard analysis profile is also shown. The simplest case is when both $T_e \sim T_i$ and $n_e \sim n_i$. With no particle or energy sources, ion-electron collisions will cause the plasma to relax to $T_e = T_i$ and this is confirmed during this pulse by measurements of T_i from the edge Charge Exchange Recombination Spectroscopy (CXRS) diagnostic standard analysis. The plasma fuel is deuterium for this pulse, so $n_i = n_e$ should be true if there are no impurities present. Measurement by the visible spectroscopy diagnostic gives the average $Z_{eff} = \sum_j n_j Z_j / n_e \approx 3$ for the H-Mode phase of this pulse which, if the impurity is entirely carbon (of which the first wall is made) implies $p/p_e \approx 1.66$. This is the lowest fraction that can be expected and the higher Z impurities that are present, the closer to 2 it should be. It should also be noted that Z_{eff} measurements are often available from CXRS and while not present for this pulse, they usually give a lower value of Z_{eff} than the visible spectroscopy.

The agreement between the MAP estimate p profile and $2p_e$ is remarkably good at the very edge and in the core. It agrees far better than the clearly over-restricted p profile of the EFITJ standard analysis which overestimates the pressure at the magnetic axis by 50%. The effect on both the flux surfaces and Q profiles is large in the plasma core which explains the difficulties analysing other diagnostics with the EFITJ flux surfaces. Using this MAP estimate instead of the EFITJ solution for the fixed flux surface n_e profile inversions of section 4.1 performs much better, implying that the flux surfaces of the CT with equilibrium prior MAP estimates, are far more accurate in the core.

6.4.2. Pedestal evolution

Figure 6.9 shows the pedestal pressure p_{ped} at $R = 3.72m$ from the MAP estimate of an inversion performed at 5ms intervals through the pulse using the simple 3-knots and a Gaussian peak model as in section 6.3.3. This is compared to the $2p_e$ of the nearest HRTS standard analysis results (red), which gives a fairly direct measurement at the relatively low temporal resolution of 20Hz. Also shown (black) is $2p_e$ using T_e taken from the standard analysis electron cyclotron emission (ECE) diagnostic and n_e taken from the the interferometry inversion as in section 4.1 based on the MAP ψ_N estimate at each time point. Both the ECE and interferometry have a temporal resolution similar to the 5ms intervals analysed here.

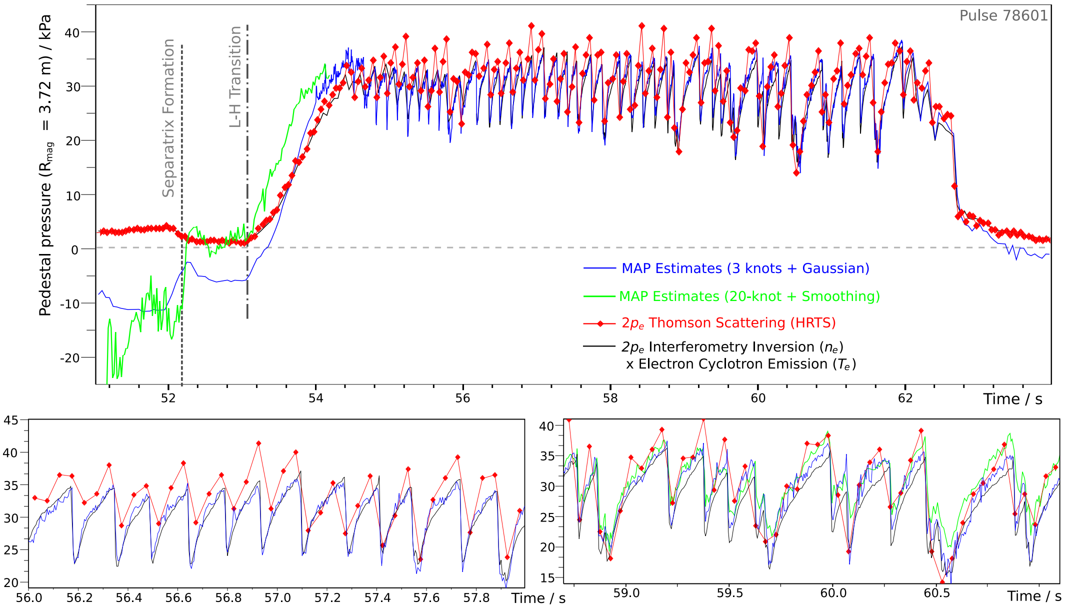

Figure 6.9.: Evolution of the pedestal pressure at $R = 3.72m$ at 5ms intervals through a type-I ELMy H-Mode pulse from MAP estimate of $P(j_\phi, p', ff' \,|\, D_m, \text{Equilibrium})$ using simple 3-knot and edge peak parametrisation (blue) for p' and ff'. For comparison, $2p_e$ derived from two independent diagnostic(s) is also shown. Two smaller time ranges are shown expanded and alternate MAP estimates using 20-knot spline and smoothing priors is shown in green.

Over the duration of the H-mode part of the pulse, the evolution of the magnetics/equilibrium inferred pedestal pressure follows the kinetic measurement well in both trend and, somewhat surprisingly given $Z_{eff} \approx 3$, in absolute magnitude. The strong parametrisation of the edge Gaussian peak makes the inversion faster and more reliable but the results are seen if a 20-knot linear interpolation profile is used with smoothing priors selected to discourage current-holes and non-monotonic pressures. The green traces show the results of inversions using the selected prior of figure 6.8. These two cases represent a significant change to the method and priors, yet roughly the same results are seen. In either case, the priors are identical for every time point so the very accurate following of the collapse and build-up between ELMs can have no source other than the magnetic measurements and proves beyond any doubt that they do provide significant accurate information from inside the plasma boundary.

Between the separatrix formation and the L-H transition, the MAP estimate is much less accurate in magnitude and shows a negative pressure for the strong parametrisation (while not physical, this was not ruled out by the priors or parametrisation). Despite this, the inferred values follow the trend of the kinetic trace after 52.2s. Before this, an abrupt change in the inferred value is seen, which coincides almost exactly with the separatrix formation, before which the plasma boundary is in direct contact with the poloidal limiter. The effect is even clearer in the more weakly parametrised MAP estimates (green). This suggests that the presence of the X-point on the boundary surface might play a special role in the way in which the magnetic diagnostics provide information about the edge currents. Very recently, this has also been identified by a more direct analytical method [87].

At each ELM, the drop in inferred pressure coincides with the drop observed by the interferometry and ECE. This shows that the response of the poloidal magnetics is at least as fast as the 5ms inversion interval and so evolution over the inter-ELM period, which is rarely shorter than 25ms for type-I ELMy H-modes at JET, can be observed clearly.

6.4.3. Parallel current evolution

The evolution of $j_{\parallel}$ vs ∇p in the pedestal is of particular interest, since it is believed that ELMs are triggered at a threshold in either $j_{\parallel}$ (Peeling mode), in ∇p (Ballooning mode) or in both (coupled Peeling-Ballooning mode). Figure 6.10 shows the flux surface average parallel and perpendicular current, integrated over the pedestal region ($0.8 \leq \psi_N < 1.0$) from each MAP estimate.

For all the MAP estimates, the equilibrium prior was well satisfied so $j_{\perp}$ relates to the pressure gradient almost exactly. The integral of $< j_{\perp} >$ over the pedestal region is therefore effectively the same as the pedestal pressure trace in figure 6.9 and is shown here only to compare against $< j_{\parallel} >$. The most striking difference is that while $< j_{\perp} >$, averaged over the ELMs (broken lines), remains relatively constant during the main H-Mode part of the pulse ($54s < t < 62s$), $< j_{\parallel} >$ slowly increases. After the start of NBI heating at 53s, $< j_{\perp} >$ rises more rapidly than $< j_{\parallel} >$ and after the heating ceases at 62.5s, $< j_{\perp} >$ falls rapidly while $< j_{\parallel} >$ decreases slowly. It is possible that this is related to the relatively long current diffusion time (of order seconds). For the equilibrium to be satisfied rapid changes in pressure gradient must be followed almost immediately by $j_{\perp}$ but $j_{\parallel}$ does not need to react immediately.

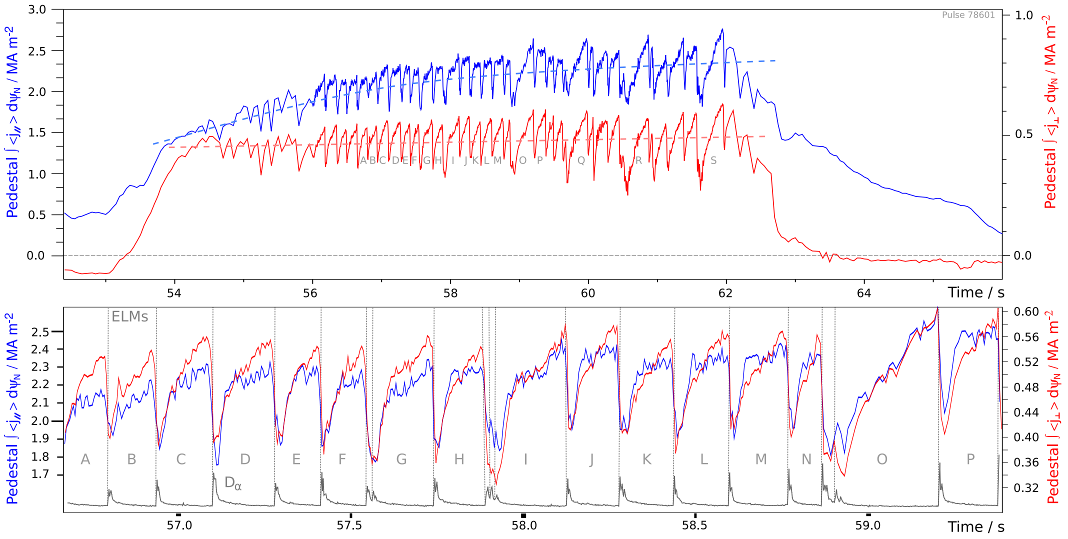

Figure 6.10.: Evolution of the flux-surface average parallel and perpendicular currents, integrated over the pedestal region ($0.8 \leq \psi_N < 1.0$) at 5ms intervals through a type-I ELMy H-Mode pulse from MAP estimate of $P(j_\phi, p, f \,|\, D_m$, Equilibrium). Bottom: Expanded plot of regular ELMs including D_α emission showing ELM crashes.

The inter-ELM detail is harder to draw even qualitative conclusions from without proper examination of the uncertainties. Given the results of the equivalent simulations in section 6.3.3, it is possible that each variable may be polluting the inference of the other. The fact that the two traces show very similar inter-ELM evolution suggests this may be the case, although it remains possible that this is a real physical effect, given that models for the parallel current usually depend on the temperature and density gradients [88]. What differences there are between the inter-ELM $< j_\parallel >$ and $< j_\perp >$ evolution does not appear to be strongly consistent. Many of the inter-ELM periods (D,H,J,K,L,M,N and P) show a initially sharp rise in $< j_\parallel >$ followed by a slower increase (sometimes almost none) while $< j_\perp >$ continues to increase all the way up to the next ELM. The longer ELM-free periods (O,Q,R and S) however, all show almost exactly the same trend in both components.

These results demonstrate the capability of the method to extract considerable detail of the physical quantities of interest from the JET magnetic data. Unfortunately, there was insufficient time within this project to conduct a detailed examination of the results.

6.5. Assessment of Bayesian Equilibrium with Polarimetry

One of the principal objectives of this work was to use the polarimetry and interferometry forward models together with magnetics and the equilibrium prior. The polarimetry should in principle dramatically improve the inferred information about the current distribution. This is especially true for the plasma edge, since the outboard most channel (channel 4, see figure 3.3) of the polarimetry effectively measures the poloidal field just inside the pedestal. The field outside the LCFS can already be accurately inferred from the magnetic diagnostics and both of these provide high enough temporal resolution to observe the inter-ELM evolution. The difference between the two relates directly to the pedestal toroidal current but the absolute uncertainty of the channel 4 measurement must be less than the difference. This is necessarily a small fraction of the magnitude, which is approximately proportional to the total plasma current. The diagnostic random noise is far below this level and while the diagnostic calibration uncertainty is possibly large in comparison, it will be approximately constant throughout the flat-top part of the plasma pulse, so the inter-ELM variation should be inferable. Unfortunately, the oscillation with line integrated density that is present in the data is larger than the inter-ELM variation and manifests clearly in the inferred pedestal current, obscuring any useful information. To illustrate this, figure 6.11 shows the standard analysis rotation signal $\Delta\psi \sim \int n_e B_Z \, dl$ for channel 4, versus time and versus line integrated density (l) for the same channel.

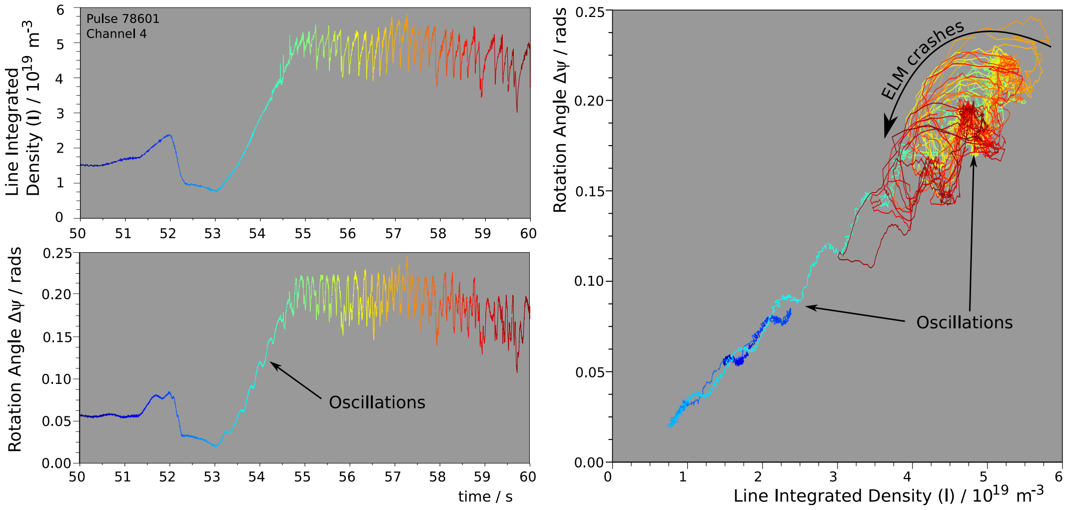

Figure 6.11.: Polarimetry rotation angle $\sim \int n_e B_Z \, dl$ for edge channel (4) versus time and line intgrated density, showing obscurity of inter-ELM data by oscillations with line integrated density.

Both the oscillations and the drop at the ELMs are clearly visible in the signal. Variations in the pedestal density should appear as a change along the main diagonal in $\psi(l)$ and changes in only the poloidal field should appear vertically (at constant l). Attempts were made to develop a post-processing method to isolate and remove the oscillation but subtle changes in phase and amplitude through the pulse make this difficult. Removal of the oscillation, either by post-processing or by isolating and removing the cause, should allow a reasonably accurate and very high temporal resolution inference of the pedestal current variation.

7. Conclusions

The objective of the wider Bayesian analysis project at JET is to improve the knowledge of the plasma state, by inferring a single consistent description with Bayesian probability theory, using forward models and data from as many as possible of the available JET diagnostics. Building on the previously developed magnetic diagnostic model, this thesis work has contributed detailed forward models for 4 JET diagnostic systems: the interferometer, polarimeter and the core and edge LIDAR Thomson Scattering systems. The models have been used, in some cases independently and sometimes in combinations, to infer information about the electron density, electron temperature and current in the JET plasma, as well as to investigate several physical effects involving these parameters.

These advanced probabalistic methods have been used to extract useful information from data that has been available for many years. For each part of this work, attempts to use the same data had been made previously with standard methods, but had either been unsuccessful or had made limited use of the information that was really available, underestimating what is possible. With the new methods, this large collection of previous plasma pulses can be re-examined in greater detail.

7.0.1. Interferometry

The Interferometry model was implemented and used with the linear Gaussian inversion to reconstruct electron density profiles with a complete description of uncertainty. It was shown that by adding reasonable prior assumptions there was sufficient information in only 8 numbers, to not only infer accurate profiles, but also to highlight problems with the magnetic geometry used in normal analyses. Coupled instead with the appropriate magnetic diagnostics and the plasma model, the interferometry was used to help infer the plasma current - something that would not normally be considered.

7.0.2. Polarimetry

A model was developed for the full evolution of the polarisation of far-infrared light through Tokamak plasmas and coupled with a model for the specifics of the JET polarimetry system. Unfortunately, complexities of the diagnostic system significantly hampered the polarimetry analysis in this work. Measured data was compared with predictions of these models, based on EFIT equilibria and density profiles from the interferometry inversion. This showed that both the instrument calibration and equilibrium reconstruction could be significantly inaccurate in medium and high density H-mode plasmas. Basing the models on current tomography instead, with a weak j_ϕ regularisation prior, showed that the predictions were strongly effected by the current distribution. While unable to help isolate the calibration issues, this indicated that the polarimetry would provide a great deal of information about j_ϕ, if the calibration issues could be resolved.

The model predictions were then used to asses two contradicting derivations, made in previous theoretical papers for the small effects of finite temperature, alongside the cold plasma approximation. The effects were smaller than the calibration uncertainties, so as well using over 15000 data points, the experience gained developing the polarimetry model was used to treat the calibration issues. The most complete model, which included relativistic effects, was shown to be in much better agreement - verifying the model for use in future Tokamak polarimetry analysis and giving the first experimental observation of relativistic finite temperature effects on plasma polarimetry.

7.0.3. LIDAR

A highly detailed forward model was developed for the JET LIDAR Thomson scattering systems and used simultaneously with both the core and edge LIDAR systems at JET. This model included the effects of the instrument function (for the first time on a LIDAR TS system), both as a convolution effect on the prediction and as a reducing and correlating effect on the photo-electron counting noise.

The model also included nuisance parameters for the many calibration factors involved, allowing the effects of their uncertainty to be automatically included in the inferred physical quantities. A number of novel methods were developed to obtain more accurate values for the calibration factors directly from the available data, with one using a special Bayesian analysis technique for handling the inference of variances and outliers.

By combining the two LIDAR systems with the interferometry, the effects of these uncertainties were reduced and accurate, high resolution electron temperature and density profiles were inferred, despite some calibration parameters remaining unknown. The remaining calibrations were then determined by performing the inversion on a large number of plasmas.

The calibration uncertainties have meant that the edge LIDAR system has been largely overlooked, due a strong belief that it's line of sight did not reach the top of the H-mode pedestal. With the combined analysis, it was immediately clear that the edge LIDAR system gives very high-resolution detail of the full pedestal in many plasma. Using this ability, the pedestal shape and approximate evolution over the ELM cycle was examined for a series of similar plasmas, showing the possibility of different phases of the pedestal recovery.

7.0.4. Bayesian Equilibrium

The uncertainty in the magnetic topology contributes much of the uncertainty in the analysis of the diagnostics, as was demonstrated for the Interferometry, LIDAR systems and especially the Polarimetry. This was true as a properly evaluated uncertainty using the current tomography (CT) model and j_ϕ regularisation prior, or as an inaccuracy in the EFIT solution.

The magnetic topology uncertainty was reduced by adding the equilibrium condition as a Bayesian prior to the CT model and is the first known attempt to treat the experimental plasma equilibrium using Bayesian analysis. The method was used to study exactly what could be inferred about the internal plasma pressure and current, from only external poloidal magnetic diagnostics. Contrary to claims of some theoretical papers, it was shown that information can be inferred about the plasma inside the boundary, especially for strongly shaped and diverted plasmas, and that this information was concentrated at the plasma edge. The relationships in β_θ, l_i and μ_i and the large uncertainty beyond these predicted for near-circular cases, was shown with a simulated L-mode limited plasma with low elongation. This was done by sampling the equilibrium posterior, effectively exploring all possible plasmas consistent with the observations and near equilibrium.

For diverted H-mode plasmas, the posterior exploration could not be performed and an iterative linear method was developed to find the maximum posterior. With this, it was shown that with the reasonable assumption of pressure monotonicity, much of the pressure profile could be retrieved from a simulated equilibrium with local structure, with especially high accuracy at the plasma edge. In a real JET H-mode pulse, good agreement was shown between the inferred pressure and the electron pressure profile p_e, taken from kinetic diagnostics. Concentrating on the plasma edge, the pressure at the top of the pedestal was examined throughout a type-I ELMy H-mode pulse, showing very good agreement in both magnitude and evolution with the kinetic p_e. The agreement suggests that the pedestal ion pressure must follow p_e closely throughout the ELM cycle and the approach effectively introduces a new and independent diagnostic of the total pedestal pressure.

The ability to infer and separate the parallel and perpendicular components of the pedestal current was examined for a simulated equilibrium, showing that information on both was present, although some contamination between the two does occur. Based on this, and the clearly well inferred perpendicular component (i.e pressure), the evolution of the parallel current was inferred for the real H-mode pulse - an important quantity that is typically difficult to obtain in Tokamak plasmas. Unfortunately, there was insufficient time to investigate the accuracy and implications of the result.

Overall, the Bayesian equilibrium method was able to infer significantly greater detail than had been seen before, or was deemed possible from the magnetic diagnostics. The scope for further development is also large, since the iterative LGI method (and software) used are not limited in resolution, other than by the number of computers available. The equilibrium priors and plasma models are entirely separate modular components, so modification by, for example, allowing for plasma flow and anisotropic pressure, will be relatively simple. The modularity also allows for almost immediate use on other Tokamaks whose magnetic diagnostics are modelled in the Minerva framework (e.g. MAST). Coupled with other diagnostic models, such as the polarimetry (assuming the calibration issues were addressed), the inferred detail is likely to be even greater. Finally, with a little further development of the exploration algorithms, and certainly with more computation power, it should be possible to sample the posterior for H-mode plasmas, giving a full description of uncertainty in the equilibrium for any combination of diagnostics, models and assumptions.

A. JET Polarimeter unknown optics model and calibration

This appendix contains, for the sake of reference, the details of the existing model that attempts to account for some unknown optical effects in the JET Polarimeter system. It adds the extra parameter D, as a non-linear part of C.

Polarisation descriptions

For directly converting between the description of polarisation as amplitude ratio/phase shift (θ,ϕ) and as principal and ellipticity angles (ψ,χ):

$$\tan 2\psi = \tan 2\theta \, \cos\phi \qquad \cos 2\theta = \cos 2\chi \, \cos 2\psi \tag{A.1}$$

$$\sin 2\chi = \sin 2\theta \, \sin\phi \qquad \tan\phi = tan 2\chi \, / \, \sin 2\psi \tag{A.2}$$

$$\tag{A.3}$$

Detected signals

The JPF nodes used are:

g4-hwp[ch]	Half Wave Plate position - gives initial polarisation (which is the plasma polarisation ψ_p during calibration).
g4-rms[ch]	Root-Mean-Square of i signal: $\langle i \cdot i \rangle$
g4-rmp[ch]	Root-Mean-square of Primed i signal: $\langle i' \cdot i' \rangle$
g4-psd[ch]	Phase-Sensitive Detected signal: $\langle i \cdot p \rangle$
g4-psp[ch]	Phase-Sensitive detected part-Primed signal: $\langle i' \cdot p \rangle$

$$R = PSD \, / \, RMS \tag{A.4}$$

$$R' = PSP \, / \, \sqrt{RMS^2 + RMP^2} \tag{A.5}$$

The following give the 'detected' polarisation in both forms to/from R,R'. This encodes the ampltiude ratio θ_D of the i and p components at the detectors but the phase difference ϕ_D between these components **before the recombination plate** that mixes the probing and frequency shifted beams. It also contains a fixed phase offset due to the electronics.

$$R = C^{-1} \tan\theta_D \, \cos\phi_D \tag{A.6}$$

$$R' = C^{-1} \tan\theta_D \, \sin\phi_D \tag{A.7}$$

$$\tan\phi_D = R \, / \, R' \tag{A.8}$$

$$\tan\theta_D = C \, R \, / \, cos\phi_D \tag{A.9}$$

$$\tan\psi_D = \frac{2\,C\,R}{(1 - C^2\,(R^2 + R'^2))} \tag{A.10}$$

$$\tan\chi_D = \frac{2\,C\,R'}{(1 + C^2\,(R^2 + R'^2))} \tag{A.11}$$

Forward function for calibration (ψ_p, $\chi_p = 0$) $\rightarrow$ (R, R')

Thses give what the detected signals (R,R') would be given the plasma polarisation (ψ_p, χ_p) and the calibration constants.$(C, D, \xi, \phi_0, \theta'_0)$. However, they are **only true when there is no plasma/initial ellipticity -** $\chi_p = 0$.

$$\theta^* = \psi_p - \theta'_0 \tag{A.12}$$

$$\begin{pmatrix} R \\ R' \end{pmatrix} = \frac{1}{C\,(1 + \cos 2\xi \cos 2\theta^*)} \begin{pmatrix} \cos\phi_0 & -\sin\phi_0 \\ \sin\phi_0 & \cos\phi_0 \end{pmatrix} \cdot \begin{pmatrix} \sin 2\theta^* \\ -\sin 2\xi \cos 2\theta^* \end{pmatrix} \tag{A.13}$$

The calibration proceedure is simply to fit these so that (R,R') matches the detected signals for the scan of ψ_p which here is ψ_p =g4-hwp[ch], by varying C, D, ξ, ϕ_0 and θ'_0.

Inversion $(R,\ R') \rightarrow (\psi_p,\ \chi_p)$

Obtaining the plasma polarisation $(\psi_p,\ \chi_p)$ from the raw signals $(R,\ R')$ and calibration constants.$(C, D, \xi, \phi_0, \theta_0')$

$$\underline{\mathrm{UR}} = \begin{pmatrix} \cos\phi_0 & \sin\phi_0 \\ -\sin\phi_0 & \cos\phi_0 \end{pmatrix} \begin{pmatrix} R \\ R' \end{pmatrix} \tag{A.14}$$

$$\underline{\mathrm{r}_{\mathrm{aux}}} = \frac{C\,\underline{\mathrm{UR}}}{1 + C\,D\,UR_1} \tag{A.15}$$

$$\underline{\mathrm{r}} = \frac{2\underline{\mathrm{r}_{\mathrm{aux}}}}{1 + r_{aux}^2} \tag{A.16}$$

$$F = \frac{(1 - r_{aux}^2)}{(1 + r_{aux}^2)} \tag{A.17}$$

$$\chi_p = \tfrac{1}{2}\sin^{-1}\left(r_2 \cos 2\xi + F \sin 2\xi\right) \tag{A.18}$$

$$\psi_p = \theta_0' + \tan^{-1}\left(F \cos 2\xi - r_2 \sin 2\xi\right) \tag{A.19}$$

Forward function for plasma $(\psi_p,\ \chi_p \neq 0) \rightarrow (R,\ R')$

This was not derived as part of the original work, it is just a direct mathematical reversal of the inversion in section A.

$$\begin{aligned} A &= \cos 2\xi\ \tan\left(2\psi_p - 2\theta_0'\right) & \text{(A.20)} \\ B &= \sin 2\xi\ \tan\left(2\psi_p - 2\theta_0'\right) & \text{(A.21)} \\ F &= \cos 2\xi\ /\ \sin 2\chi_p & \text{(A.22)} \\ G &= \sin 2\xi\ /\ \sin 2\chi_p & \text{(A.23)} \end{aligned}$$

$$
\begin{aligned}
a &= \frac{A^2F^2}{(1+G)} + B^2(1+G) + 2ABF + (1+G) && \text{(A.24)}\\
b &= -2F - \frac{2A^2F}{(1+G)} - 2AB && \text{(A.25)}\\
c &= \frac{A^2}{(1+G)} + (1-G) && \text{(A.26)}\\
r_{aux_2} &= (-b \pm \sqrt{b^2-4ac})/2a && \text{(A.27)}\\
& && \text{(A.28)}
\end{aligned}
$$

Pick the smallest real solution of the quadratic (r_{aux_2}).

$$
r_{aux_1} = -\left(\frac{AC}{(1+D)} + B\right) r_{aux_2} + \frac{A}{(1+D)}; \tag{A.29}
$$

$$
C \begin{pmatrix} R \\ R' \end{pmatrix} = \begin{pmatrix} \cos\phi_0 & -\sin\phi_0 \\ \sin\phi_0 & \cos\phi_0 \end{pmatrix} \underline{\mathrm{r}}_{aux} \tag{A.30}
$$

Cold plasma line-integral approximations

The cold-plasma approximations for seperated Faraday/Cotton-Mouton effects are:

$$
\begin{aligned}
\Delta\psi &\approx \frac{e^3}{\omega^2\epsilon_0 m_e^2 c} \int n_e(z)\, B_z(z)\, dz && \text{(A.31)}\\
\chi &\approx \frac{e^4}{2\omega^3\epsilon_0 m_e^3 c} \int n_e(z)\, B_\perp^2(z)\, dz && \text{(A.32)}
\end{aligned}
$$

B. Magnetic Diagnostic Coils Rejection Procedure

The JET poloidal magnetic diagnostics presently consist of 230 pickup coils, 88 saddle coils and 6 full toroidal flux loops. For each pulse and/or time point any coil may suffer complete failure or large systematic error beyond their random noise. Each day, JET performs a *dry-run* pulse where each PF coil is turned on individually. No plasma is present so the poloidal field is created only by the PF coils and possibly iron core eddy currents. The following admittedly ad-hoc procedure is used in this work to reject suspect coils:

1. Enable all diagnostic coils.
2. Load the data for the nearest dry-run pulse to the plasma pulse desired.
3. Estimate the random noise level σ from the scatter on the baseline before poloidal field is present.
4. For time slices every 500ms including the switching of every PF coil, repeat the following until the worst fitting diagnostic coil has $\chi^2 < 4$:
    a) Using $j_\phi = 0$ for the plasma, fit the iron core currents to best match the predictions to data of the remaining diagnostic coils.
    b) Reject the coil with the highest χ^2.
5. Enable only coils which had $\chi^2 < 4$ for all time slices of the dry-run.
6. Load the data of the plasma pulse and time slice of interest.
7. Repeat the following until every diagnostic coil has $\chi^2 < 10$:
    a) Using a very weak CAR prior perform a standard CT LGI for both plasma and iron core currents.
    b) Reject the coil with the highest χ^2.

Figure B.1 shows the 4 saddle and 29 pick-up coils rejected during the dry-run test and the 4 further pick-up coils rejected during the plasma test for the pulse and time point from which most of the results in section 6.4 are derived.

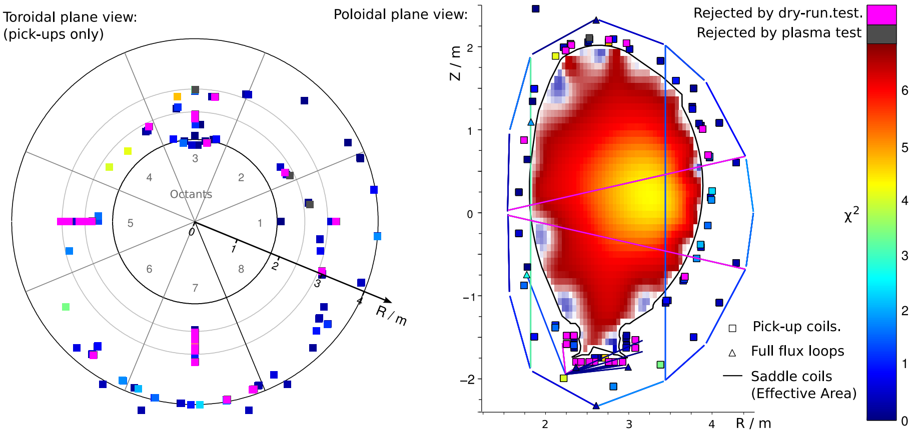

Figure B.1.: Magnetic diagnostic coils rejected during dry-run cross check with PF coils (magenta) and during basic CT test on a typical time point (grey). Also shown are is the MAP toroidal plasma currents at the final inversion and the χ^2 of the remaining coils.

NB: Not all coils are visible as many share toroidal/poloidal coordinates with others. Rejected coils are shown on top.

C. Glossary - Terms and Acronyms

This section gives, for ease of reference, terms and acronyms used in this document. The 'scope' indicates whether the term is used only in this document (Document), in the Bayesian and/or integrated analysis communities (Analysis), at JET (JET) or throughout the Tokamak community (Tokamak).

Term/Acronym	Brief description	Scope	Page
CT	Current Tomography	Document	52
LGI	Linear Gaussian Inversion	Document	44
TLGI	Truncated Linear Gaussian Inversion	Document	44
MCMC	Markov Chain Monte Carlo (Typically Metropolis-Hastings Algorithm)	Analysis	50
GA	Genetic Algorithms	Analysis	49
Minerva	Software and conceptual framework in which this work was carried out.	Document	42
MAP	MAximum Posterior	Analysis	42
Forward Model	Diagnostic model that gives most likely data for given physics and calibration state.	Analysis	38
JET	Joint European Torus - The Tokamak this work covers	Tokamak	28
MAST	Mega-Amp Spherical Tokamak - A spherical tokamak also based on the Culham site with JET.	Tokamak	43

TF	Toroidal field (usually coils)	Tokamak	16
PF	Poloidal field (usually coils)	Tokamak	17
Vacuum Field	Magnetic field without the part created by the plasma currents	Tokamak	17
Ripple	Variation of TF due to finite number of TF coils.	Tokamak	16
Dry-Runs	Full pulses with no plasma for magnetics calibration. Also useful for some diagnostic calibration.	JET	170
First Wall	Geometry of last material surface facing the plasma.	Tokamak	17
LCFS	Last Closed Flux Surface	Tokamak	17
Magnetic Axis	Centre of nested flux surfaces in poloidal plane.	Tokamak	17
Normalised Flux	Magnetic flux (usually poloidal) normalised to 0 at magnetic axis and 1 at LCFS	Tokamak	20
X-point	Point of zero poloidal field, usually on the separatrix.	Tokamak	18
Separatrix	Flux surface that passes through the X-point in a diverted geometry plasma.	Tokamak	18
SOL	Scrape Off Layer	Tokamak	18
Divertor	Special part of vessel/first wall used extract exhaust.	Tokamak	18
Safety Factor q	Related to average magnetic field line pitch angle of a flux surface	Tokamak Tokamak	17
Beta Poloidal (β_θ)	Ratio of volume integral of pressure to poloidal field integrated along LCFS.	Tokamak	22
Strike Points	Contact points of separatrix on the first wall.	Tokamak	18
H-Mode	High confinement mode due to an ETB.	Tokamak	24
L-Mode	Low confinement mode.	Tokamak	24
Pedestal	Large gradient region of equilibrium profiles (n_e, T_e etc) at plasma edge.	Tokamak	25
ETB	Edge Transport Barrier	Tokamak	24
ELM	Edge Localised Model	Tokamak	26
Ohmic Plasma	Plasma driven only inductively by main transformer.	Tokamak	28
NBI	Neutral Beam Injection	Tokamak	28
ICRH	Ion Cyclotron Resonant Heating	Tokamak	28
EFIT(J)	Equilibrium FITting Code (JET)	Tokamak	21
Standard Analysis	JET automatic standard data processing (A.K.A 'Chain 1')	JET	30

TS	Thomson Scattering	Tokamak	98
LIDAR	Light Detection and Ranging (Special case TS system)	JET	102
PMT	Photo Multiplier Tube	Diagnostics	107
ADC	Analog to Digital Converter	Diagnostics	107
Ambient Light	LIDAR-TS: Constant background plasma light not from TS.	Document	108
Stray Light	LIDAR-TS: Short light spikes/pulses possibly from reflections etc	Document	114

Bibliography

[1] O. Ford et al. 'Forward modeling of JET polarimetry diagnostic'. *Review of Scientific Instruments*, **79**, 10:10F324, 2008. doi: 10.1063/1.2956880.
`http://link.aip.org/link/?RSI/79/10F324/1.`

[2] O. P. Ford et al. 'Experimental verification of relativistic finite temperature polarimetry effects at JET'. *Plasma Physics and Controlled Fusion*, **51**, 6:065004, 2009. doi: 10.1088/0741-3335/51/6/065004.
`http://stacks.iop.org/0741-3335/51/i=6/a=065004.`

[3] O. P. Ford et al. 'Bayesian combined analysis of JET LIDAR, edge LIDAR and interferometry diagnostics'. In 'Proc. of the 36th EPS Conference on Plasma Physics, Sofia, Bulgaria', P2.150. 2009.
`http://epsppd.epfl.ch/Sofia/pdf/P2_150.pdf.`

[4] J. Svensson et al. 'Connecting physics models and diagnostic data using bayesian graphical models'. In 'Proc. of the 37th EPS Conference on Plasma Physics, Dublin, Ireland', O4.117. 2010.
`http://ocs.ciemat.es/EPS2010PAP/pdf/O4.117.pdf.`

[5] J. Lawson. 'Some criteria for a power producing thermonuclear reactor'. *Proceedings of the Physical Society. Section B*, **70**, 1:6, 1957.
`http://stacks.iop.org/0370-1301/70/i=1/a=303.`

[6] J. Wesson. *Tokamaks (Second Edition)*. Oxford University Press, 1997. ISBN 0-19-856293-4X.

[7] A. Thyagaraja and K. G. McClements. 'Toroidal and poloidal flows in single-fluid and two-fluid tokamak equilibria'. *Physics of Plasmas*, **13**, 6:062502, 2006. doi: 10.1063/1.2205189.
`http://link.aip.org/link/?PHP/13/062502/1.`

[8] M. c HA. *Plasma rotation in the MAST and JET tokamaks*. Ph.D. thesis, 2009.
`http://www.minhduchua.free.fr/fusion/MDH_thesis_soft_copy.pdf.`

[9] K. Cromb. 'Poloidal rotation velocity in JET advanced mode plasmas using charge exchange recombination spectroscopy'. In 'Proc. of the 33rd EPS Conference on Plasma Physics, Rome, Italy', D5.015. 2001.
`http://epsppd.epfl.ch/Roma/pdf/D5_015.pdf.`

[10] V. Shafranov. 'On magnetohydrodynamical equilibrium configurations.' *Zhurnal Experimentalnoi i Teoretiches-koi Fiziki*, **33**, 3:710–722, 1957.

[11] R. Lust and A. Schluter. 'Axialsymmetrische magnetohydrodynamische gleichgewichtskonfigurationen.' *Zietschrift fur Naturforschung*, **12A**, 10:850, 1957.

[12] H. Grad and H. Rubin. 'Hydromagnetic equilibria and force-free fields'. In 'Proc. 2nd United National Intl. Conf. on Peacful Uses of Atmoic Energy, Geneva', volume 31, page 190. 1958.

[13] B. J. Braams. 'The interpretation of tokamak magnetic diagnostics'. *Plasma Physics and Controlled Fusion*, **33**, 7:715–748, 1991. doi: 10.1088/0741-3335/33/7/001.
`http://stacks.iop.org/0741-3335/33/i=7/a=001.`

[14] D. O'Brien et al. 'Equilibrium analysis of iron core tokamaks using a full domain method'. *Nuclear Fusion*, **32**, 8:1351, 1992.
`http://stacks.iop.org/0029-5515/32/i=8/a=I05.`

[15] V. D. Shafranov. 'Determination of the parameters β_i and l_i in a tokamak for arbitrary shape of plasma pinch cross-section'. *Plasma Physics*, **13**, 9:757, 1971.
`http://stacks.iop.org/0032-1028/13/i=9/a=006.`

[16] V. D. Pustovitov. 'On the theory of magnetic diagnostics in stellarators and tokamaks'. *Fusion Eng. Des.*, **34-35**:689–693, 1997.
`http://dx.doi.org/10.1016/S0920-3796(96)00565-0.`

[17] J. P. Freidberg et al. 'Why β_{pol} and l_i cannot be separately measured in a near circular tokamak'. *Plasma Physics and Controlled Fusion*, **35**, 11:1641–1648, 1993. doi: 10.1088/0741-3335/35/11/010.
`http://stacks.iop.org/0741-3335/35/i=11/a=010.`

[18] L. E. Zakharov et al. 'The theory of variances in equilibrium reconstruction'. *Physics of Plasmas*, **15**, 9:092503, 2008. doi: 10.1063/1.2977480.
`http://link.aip.org/link/?PHP/15/092503/1.`

[19] L. L. Lao et al. 'Separation of β_p and l_i in tokamaks of non-circular cross-section.' *Nucl. Fus.*, **25**, 10:1421–1436, 1985.

[20] F. W. annd others. 'Regime of improved confinement and high beta in neutral-beam-heated divertor discharges of the ASDEX tokamak'. *Physical Review Letters*, **49**, 19:1408–1412, 1982.

[21] F. Wagner et al. 'Development of an edge transport barrier at the H-mode transition of ASDEX'. *Physical Review Letters*, **53**, 15:1453–1456, 1984.

[22] B. Coppi et al. 'Neoclassical transport and the principle of profile consistency'. *Comments on Plasma Physics and Controlled Fusion*, **5**, 6:261–270, 1980.

[23] X. Garbet et al. 'Profile stiffness and global confinement'. *Plasma Physics and Controlled Fusion*, **46**, 9:1351, 2004. doi: 10.1063/1.873240.
`http://stacks.iop.org/0741-3335/46/i=9/a=002.`

[24] E. J. Doyle et al. 'Modifications in turbulence and edge electric fields at the L-H transition in the DIII-D tokamak'. *Physics of Fluids B: Plasma Physics*, **3**, 8:2300–2307, 1991. doi: 10.1063/1.859597.
`http://link.aip.org/link/?PFB/3/2300/1.`

[25] M. Brix et al. 'Accuracy of efit equilibrium reconstruction with internal diagnostic information at jet'. *Review of Scientific Instruments*, **79**, 10:10F325, 2008. doi: 10.1063/1.2964180.
`http://link.aip.org/link/?RSI/79/10F325/1.`

[26] R. Pasqualotto et al. 'High resolution thomson scattering for JET'. *Review of Scientific Instruments*, **75**, 10:3891–3893, 2004. doi: 10.1063/1.1787922.
`http://link.aip.org/link/?RSI/75/3891/1.`

[27] D. V. Bartlett et al. 'Overview of JET ECE measurements.' In 'Proc. of the 6th Joint Workshop on ECE and ECRH, Oxford', 1987.

[28] D. V. Bartlett et al. 'Recent progress in the measurement and analysis of ECE on JET'. In 'Proc. of the 9th Joint Workshop on ECE and ECRH, Borrego Springs', 1995.

[29] E. de la Luna et al. 'Electron cyclotron emission radiometer upgrade on the JET tokamak'. *Review of Scientific Instruments*, **75**, 10:3831–3833, 2004. doi: 10.1063/1.1781376.
`http://link.aip.org/link/?RSI/75/3831/1.`

[30] C. R. Negus et al. 'Enhanced core charge exchange recombination spectroscopy system on JET'. *Review of Scientific Instruments*, **77**, 10:10F102, 2006. doi: 10.1063/1.2222170.
`http://link.aip.org/link/?RSI/77/10F102/1`.

[31] Y. Andrew et al. 'Improved charge exchange spectroscopy on the JET for ion temperature and rotation velocity profiles'. *Review of Scientific Instruments*, **77**, 10:10E913, 2006. doi: 10.1063/1.2228810.
`http://link.aip.org/link/?RSI/77/10E913/1`.

[32] J. Svensson and A. Werner. 'Large scale bayesian data analysis for nuclear fusion experiments'. In 'Proc. IEEE Workshop on Intelligent Signal Processing WISP', pages 1–6. 2007. doi: 10.1109/WISP.2007.4447579.

[33] A. P. Millar et al. 'Genetic algorithms in plasma diagnostic analysis'. *Plasma Physics and Controlled Fusion*, **42**, 3:337, 2000. doi: 10.1088/0741-3335/42/3/310.
`http://stacks.iop.org/0741-3335/42/i=3/a=310`.

[34] R. Fischer et al. 'Bayesian modelling of fusion diagnostics'. *Plasma Physics and Controlled Fusion*, **45**, 7:1095, 2003. doi: 10.1088/0741-3335/45/7/304.
`http://stacks.iop.org/0741-3335/45/i=7/a=304`.

[35] A. Dinklage et al. 'Topics and methods for data validation by means of bayesian probability theory'. *Fusion Science and Technology*, **46**, 2:355–364, 2004.
`http://www.new.ans.org/pubs/journals/fst/a_575`.

[36] M. Reginatto and A. Zimbal. 'Bayesian and maximum entropy methods for fusion diagnostic measurements with compact neutron spectrometers'. *Review of Scientific Instruments*, **79**, 2:023505, 2008. doi: 10.1063/1.2841695.
`http://link.aip.org/link/?RSI/79/023505/1`.

[37] J. Svensson et al. 'Integrating diagnostic data analysis for W7-AS using bayesian graphical models'. *Review of Scientific Instruments*, **75**, 10:4219–4221, 2004. doi: 10.1063/1.1789611.
`http://link.aip.org/link/?RSI/75/4219/1`.

[38] J. Svensson and A. Werner. 'Current tomography for axisymmetric plasmas'. *Plasma Physics and Controlled Fusion*, **50**, 8:085002, 2008.
`http://stacks.iop.org/0741-3335/50/i=8/a=085002`.

[39] L. C. Appel et al. 'Bayesian inference applied to magnetic equilibrium on MAST'. In 'Proc. of the 37th EPS Conference on Plasma Physics, Dublin, Ireland', P4.103. 2010.
`http://ocs.ciemat.es/EPS2010PAP/pdf/P4.103.pdf`.

[40] M. J. Hole et al. 'Model data fusion: developing bayesian inversion to constrain equilibrium and mode structure'. *Journal of Plasma and Fusion Research*, **9**:479–486, 2010.
`http://arxiv.org/abs/1002.3189`.

[41] D. S. Sivia. *Data Analysis, A Bayesian Tutorial.* Oxford University Press, 1999. ISBN 0-19-851889-7.

[42] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation.* SIAM, 2005.

[43] J. J. Dongarra and D. W. Walker. 'Software libraries for linear algebra computations on high performance computers'. *SIAM Review*, **37**, 2:151–180, 1995. ISSN 0036-1445. doi: 10.1137/1037042.
`http://dx.doi.org/10.1137/1037042`.

[44] L. S. Blackford et al. *ScaLAPACK Users' Guide.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997. ISBN 0-89871-397-8 (paperback).
`http://www.netlib.org/scalapack/slug/`.

[45] S. Geman and D. Geman. 'Stochastic relaxation, gibbs distributions, and the bayesian restoration of images'. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **PAMI-6**, 6:721 –741, 1984. ISSN 0162-8828. doi: 10.1109/TPAMI.1984.4767596.
`http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4767596.`

[46] 'Pseudo random variable generation'. Technical report, University of Chicago, Econ 319, 2002.
`http://athens.src.uchicago.edu/jenni/econ319_2003/042202_TA_sess.pdf.`

[47] J. Geweke. 'Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities'. In 'Proc. of Computing Science and Statistics: the Twenty-Third Symposium on the Interface, Seattle', 1991.
`http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.26.6892.`

[48] J. R. Shewchuk. 'An introduction to the conjugate gradient method without the agonizing pain', 1994.
`http://www.cs.cmu.edu/~jrs/jrspapers.html.`

[49] W. K. Hastings. 'Monte carlo sampling methods using markov chains and their applications.' *Biometrika*, **57**, 1:97–109, 1970. doi: 10.1093/biomet/57.1.97.

[50] G. O. Roberts and J. S. Rosenthal. 'Optimal scaling for various metropolishastings algorithms.' *Statistical Science*, **16**, 4:351–367, 2001. doi: 10.1214/ss/1015346320.

[51] C. Andrieu and J. Thoms. 'A tutorial on adaptive mcmc'. *Statistics and Computing*, **18**, 4:343373, 2008. doi: 10.1007/s11222-008-9110-y.

[52] T. H. Stix. *Waves in Plasmas.* Springer, 1992.

[53] S. E. Segre. 'A review of plasma polarimetry - theory and methods'. *Plasma Physics and Controlled Fusion*, **41**, 2:R57–R100, 1999. doi: 10.1088/0741-3335/41/2/001.
`http://stacks.iop.org/0741-3335/41/i=2/a=001.`

[54] F. P. Orsitto et al. 'Modelling of polarimetry measurements at JET'. *Review of Scientific Instruments*, **50**, 11:115009, 2008. doi: 10.1088/0741-3335/50/11/115009.
`http://stacks.iop.org/0741-3335/50/115009.`

[55] M. Brombin et al. 'Real-time electron density measurements from Cotton-Mouton effect in JET machine'. *Review of Scientific Instruments*, **79**:10E718, 2008. doi: 10.1063/1.2956831.
`http://link.aip.org/link/RSINAK/v79/i10/p10E718/s1.`

[56] K. Guenther. 'Approximate method to extract the pure faraday and cotton-mouton effects from polarimetry measurements in a tokamak'. *Plasma Physics and Controlled Fusion*, **46**, 9:1423–1441, 2004. doi: 10.1088/0741-3335/46/9/006.
`http://stacks.iop.org/0741-3335/46/1423.`

[57] C. Mazzotta et al. 'Models comparison for JET polarimeter data'. In G. Gorini, F. P. Orsitto, E. Sindoni, et al., editors, 'Burning Plasma Diagnostics: An International Conference', volume 988, pages 140–143. AIP, 2008. doi: 10.1063/1.2905056.

[58] G. Braithwaite et al. 'JET Report ET-IR(85)08)'. Technical report, Joint European Torus (JET), 1985.

[59] G. Braithwaite et al. 'JET polari-interferometer'. *Review of Scientific Instruments*, **60**, 9:2825–2834, 1989. doi: 10.1063/1.1140666.
`http://link.aip.org/link/?RSI/60/2825/1.`

[60] K. Guenther. ''complete' far-infrared polarimetry measurements at JET'. In 'Proc. of 31st EPS Conference on Plasma Physics London, UK', P5.172. 2004.
`http://epsppd.epfl.ch/London/pdf/P5_172.pdf.`

[61] K. Guenther. 'Polarimetry at JET a step by step study'. Technical report, Joint European Torus JET, Abingdon, UK., 2004.

[62] K. Guenther. 'A model accounting for the spurious ellipticity generated by the JET polarimeter'. Technical report, Joint European Torus, Abingdon, UK., 2005.

[63] K. Guenther. 'Evaluation of the polarization state of the probing beam'. Technical report, Joint European Torus, Abingdon, UK., 2005.

[64] S. E. Segre and V. Zanza. 'Finite electron temperature effects in plasma polarimetry and interferometry'. *Physics of Plasmas*, **9**, 7:2919–2925, 2002. doi: 10.1063/1.1485075.
`http://link.aip.org/link/?PHP/9/2919/1.`

[65] V. V. Mirnov et al. 'Finite electron temperature effects on interferometric and polarimetric measurements in fusion plasmas'. *Physics of Plasmas*, **14**, 10:102105, 2007. doi: 10.1063/1.2790886.
`http://link.aip.org/link/?PHP/14/102105/1.`

[66] H. S. others. 'The LIDAR Thomson scattering diagnostic on JET (invited)'. *Review of Scientific Instruments*, **59**, 8:1451–1456, 1988. doi: 10.1063/1.1139686.
`http://link.aip.org/link/?RSI/59/1451/1.`

[67] M. Kempenaars et al. 'Enhancement of the JET edge LIDAR Thomson scattering diagnostic with ultrafast detectors'. *Review of Scientific Instruments*, **79**, 10:10E728, 2008. doi: 10.1063/1.2969078.
`http://link.aip.org/link/?RSI/79/10E728/1.`

[68] I. H. Hutchinson. *Principles of Plasma Diagnostics*, chapter 7. Cambridge University Press, 1987.

[69] O. Naito et al. 'Analytic formula for fully relativistic thomson scattering spectrum'. *Physics of Fluids B: Plasma Physics*, **5**, 11:4256–4258, 1993. doi: 10.1063/1.860593.
`http://link.aip.org/link/?PFB/5/4256/1.`

[70] R. Scannell et al. 'Absolute calibration of LIDAR thomson scattering systems by rotational Raman scattering'. *Review of Scientific Instruments*, **81**, 4:045107, 2010. doi: 10.1063/1.3374111.
`http://link.aip.org/link/?RSI/81/045107/1.`

[71] A. Gelman et al. *Bayesian Data Analysis*, chapter 3.2. Chapman and Hall, 1995.

[72] R. J. Groebner et al. 'Progress in quantifying the edge physics of the H-mode regime in DIII-D'. *Nuclear Fusion*, **41**, 12:1789–1802, 2001. doi: 10.1088/0029-5515/41/12/306.
`http://stacks.iop.org/0029-5515/41/i=12/a=306.`

[73] R. J. Groebner et al. 'Role of neutrals in density pedestal formation in DIII-D'. *Plasma Physics and Controlled Fusion*, **44**, 5A:A265–A272, 2002. doi: 10.1088/0741-3335/44/5A/326.
`http://stacks.iop.org/0741-3335/44/i=5A/a=326.`

[74] J. Flanagan et al. 'Alternative calibration techniques for thomson scattering diagnostics on large fusion devices via an integrated data approach'. In 'Poster, 37th Institute of Physics Plasma Conference', 2010.

[75] M. N. A. Beurskens et al. 'Pedestal and scrape-off layer dynamics in ELMy H-mode plasmas in JET'. *Nuclear Fusion*, **49**, 12:125006, 2009. doi: 10.1088/0029-5515/49/12/125006.
`http://stacks.iop.org/0029-5515/49/i=12/a=125006.`

[76] M. N. A. Beurskens et al. 'Pedestal width and ELM size identity studies in JET and DIII-D; implications for ITER'. *Plasma Physics and Controlled Fusion*, **51**, 12:124051, 2009. doi: 10.1088/0741-3335/51/12/124051.
`http://stacks.iop.org/0741-3335/51/i=12/a=124051.`

[77] L. Garzotti et al. ‘Investigating pellet ELM triggering physics using the new small size pellet launcher at JET’. In ‘Proc. of the 37th EPS Conference on Plasma Physics, Dublin, Ireland’, P2.131. 2010.
`http://ocs.ciemat.es/EPS2010PAP/pdf/P2.131.pdf.`

[78] B. Alper et al. ‘Insight from fast data on pellet ELM pacing at jET’. In ‘Proc. of the 37th EPS Conference on Plasma Physics, Dublin, Ireland’, P2.173. 2010.
`http://ocs.ciemat.es/EPS2010PAP/pdf/P2.173.pdf.`

[79] A. Kirk, T. O’Gorman, S. Saarelma, et al. ‘A comparison of h-mode pedestal characteristics in mast as a function of magnetic configuration and elm type’. *Plasma Physics and Controlled Fusion*, **51**, 6:065016, 2009. doi: 10.1088/0741-3335/51/6/065016.
`http://stacks.iop.org/0741-3335/51/i=6/a=065016.`

[80] A. V. Chankin et al. ‘Parallel currents in the scrape-off layer of JET diverted discharges’. *Journal of Nuclear Materials*, **196-198**:739 – 744, 1992. ISSN 0022-3115. doi: 10.1016/S0022-3115(06)80134-0.
`http://www.sciencedirect.com/science/article/B6TXN-4PRFR1J-4R/2/f90807a0724c2e2dd2245a2fa09ceb25.`
Plasma-Surface Interactions in Controlled Fusion Devices, Proceedings of the Tenth International Conference on Plasma-Surface Interactions in Controlled Fusion Devices.

[81] C. S. Pitcher and P. C. Stangeby. ‘Experimental divertor physics’. *Plasma Physics and Controlled Fusion*, **39**, 6:779, 1997. doi: 10.1088/0741-3335/39/6/001.
`http://stacks.iop.org/0741-3335/39/i=6/a=001.`

[82] H. Wilson. ‘Edge localized modes in tokamaks’. *Fusion Science and Technology*, **53**, 2T:161–169, 2008.
`http://www.new.ans.org/pubs/journals/fst/a_1702.`

[83] M. F. M. D. Bock et al. ‘Edge current measurements using MSE during MAST H-modes’. In ‘Proc. 37th EPS Conference on Plasma Phys, Dublin, Ireland.’, O2.107. 2010.
`http://ocs.ciemat.es/EPS2010PAP/pdf/O2.107.pdf.`

[84] D. M. Thomas et al. ‘Measurement of pressure-gradient-driven currents in tokamak edge plasmas’. *Physical Review Letters*, **93**, 6:065003, 2004.

[85] R. G. L. Vann et al. ‘Using bernstein wave emission to measure the current in the tokamak edge’. In ‘Proc. 37th EPS Conference on Plasma Phys, Dublin, Ireland.’, P5.138. 2010.
`http://ocs.ciemat.es/EPS2010PAP/pdf/P5.138.pdf.`

[86] N. C. Hawkes et al. ‘Observation of zero current density in the core of JET discharges with lower hybrid heating and current drive’. *Physical Review Letters*, **87**, 11:115001, 2001. doi: 10.1103/PhysRevLett.87.115001.
`http://link.aps.org/doi/10.1103/PhysRevLett.87.115001.`

[87] P. McCarthy et al. ‘Recovery of edge moments of the current density profile in a tokamak from external magnetic measurements’. In ‘Proc. 37th EPS Conference on Plasma Phys, Dublin, Ireland.’, O4.116. 2010.
`http://ocs.ciemat.es/EPS2010PAP/pdf/O4.116.pdf.`

[88] O. Sauter et al. ‘Neoclassical conductivity and boostrap current formulas for general axisymmetric equilibria and arbitrary collisionality regime’. *Physics of Plasmas*, **6**, 7:2834–2839, 1999. doi: 10.1063/1.873240.
`http://link.aip.org/link/?PHP/6/2834/1.`