

Bayesiaanse geïntegreerde bepaling van de effectieve ionaire lading
via remstralings- en ladingsuitwisselingsspectroscopie
in tokamakplasma's

Bayesian Integrated Determination of the Ion Effective Charge
via Bremsstrahlung and Charge Exchange Spectroscopy
in Tokamak Plasmas

Geert Verdoolaege

Promotor: prof. dr. ir. G. Van Oost
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Toegepaste Natuurkunde

Vakgroep Toegepaste Fysica
Voorzitter: prof. dr. ir. C. Leys
Faculteit Ingenieurswetenschappen
Academiejaar 2006 - 2007



ISBN-10 90-8578-118-3
ISBN-13 978-90-8578-118-9
NUR 919, 926
Wettelijk depot: D/2006/10.500/76

Promotor

prof. dr. ir. G. Van Oost Universiteit Gent

Examencommissie

prof. dr. ir. L. Taerwe	Universiteit Gent
prof. dr. ir. C. Leys	Universiteit Gent
dr. ir. K. Crombé	Universiteit Gent
prof. dr. ir. J.-M. Noterdaeme	Universiteit Gent
dr. G. Telesca	Institut für Plasmaphysik, Forschungszentrum Jülich
dr. M. von Hellermann	FOM-Institute for Plasma Physics 'Rijnhuizen'
prof. dr. em. W. Wieme	Universiteit Gent

Universiteit Gent
Faculteit Ingenieurswetenschappen

Vakgroep Toegepaste Fysica
Jozef Plateastraat 22
B-9000 Gent, België

Tel.: +32-9-264.38.31
Fax: +32-9-264.41.98
e-mail: geert.verdoolaege@ugent.be



Dit werk werd gedeeltelijk gefinancierd door het
Fonds Wetenschappelijk Onderzoek – Vlaanderen
(projectnummers 31517102 en 3G019103)

De auteur geeft de toelating dit proefschrift voor consultatie beschikbaar te stellen en delen van dit proefschrift te kopiëren voor persoonlijk gebruik.

Elk ander gebruik valt onder de beperking van het auteursrecht, in het bijzonder tot de verplichting de bron uitdrukkelijk te vermelden bij het aanhalen van resultaten van dit proefschrift.

Geert Verdoolaege
Gent, november 2006

Monte-Carlo



Aan mijn ouders

Dankwoord

—Word of thanks—

Gedurende de voorbije jaren waarin ik onderzoek heb gevoerd aan de Universiteit Gent heb ik op verschillende vlakken, en van verschillende mensen, hulp gekregen. Op deze plaats zou ik die mensen hiervoor willen bedanken. Ik heb getracht daarbij niemand te vergeten, maar mocht dat toch het geval zijn, dan wil ik hierbij ook de gedupeerden bedanken.

Om te beginnen ben ik mijn promotor, professor Guido Van Oost, dankbaar om verschillende redenen. Ten eerste heeft hij op een zeer doeltreffende manier gereageerd op mijn blijk van interesse voor het fusieonderzoek na mijn eerste bezoek aan de TEXTOR-tokamak in Jülich. In feite was het de bedoeling dat ik bij de onderzoeksgroep Plasmatechnologie zou werken, maar het was toen de gewoonte dat elke nieuwkomer toch eens de wonderen ging aanschouwen van het fusieonderzoek in Jülich. TEXTOR is een vrij indrukwekkende installatie en ondermeer dat heeft tot mijn verbeelding gesproken¹. In het verdere verloop van mijn onderzoek steunde Guido mij ten volle en bovendien wist hij mij telkens weer in contact te brengen met andere onderzoekers die een interessante bijdrage konden leveren tot mijn werk. Maar ik ben mijn promotor vooral dankbaar voor de grote vrijheid die hij mij gelaten heeft in mijn onderzoek, wat er mede toe geleid heeft dat ik de interessante en veelbelovende weg heb kunnen inslaan van de Bayesiaanse gegevensverwerking voor fusiediagnostieken.

Ik ben ook de andere leden van de Vakgroep Toegepaste Fysica dankbaar voor de goede samenwerking gedurende de voorbije jaren. Om te beginnen is dat professor Wieme, die mij de kans bood als assistent te beginnen aan de vakgroep. Ik denk ook vooral aan Daniël Lauwers, Dries Vincke en Peter Guns voor de technische hulp die ze mij boden met mijn Z_{eff} -diagnostiek aan TEXTOR. Daniël kwam met ver-

¹In feite is dit dus niet alleen een dankwoord, maar ook een verontschuldiging, want hierbij moet ik mijn excuses aanbieden aan Christophe Leys.

schillende goede ideeën (ook op enkele kritieke momenten) in verband met de mechanische constructie van een aantal onderdelen. Daarnaast was hij ook een sturende factor bij enkele bezoeken aan Jülich en omstreken. Peter Guns heeft mij geholpen met een aantal elektronische onderdelen voor mijn diagnostiek, en de digitale timing-module is het resultaat van zijn grote vaardigheid met ook digitale componenten. Verder dank ik Eric Desoppere voor verschillende hulpzame discussies, en de collega's samen met wie ik talloze practica en oefeningenlessen gegeven heb (Michael, Frank, Peter P., ...). Ik heb ook vele aangename momenten beleefd met andere vakgroepleden, zoals tijdens de vakgroepuitstappen, de occasionele drinks en natuurlijk met het groepje waarmee we 's middags gingen eten (Christophe, Carlos, ...).

Ik dank ook de leden van het Laboratorium voor Plasmafysica van de Koninklijke Militaire School voor de goede samenwerking. In het bijzonder dank ik professor R. Weynants voor de regelingen met betrekking tot de EURATOM-mobiliteit.

De vele culturele momenten in Duitsland met Peter Peleman, Kristel Crombé en Stefan Jachmich verdienen een aparte alinea. Om te beginnen wil ik Peter bedanken voor de talloze onvermoeibare uurtjes rijden van en naar Jülich, en 's avonds naar Aken, Keulen of Düsseldorf. De discussies bij een typisch Duits (of Zuid-Amerikaans) drankje stelden dikwijls het werk even in een ander perspectief. En het eten van een braadworst op een Duitse kerstmarkt doet een mens toch beseffen dat hij deel is van een grotere wereld.

Coming to Jülich, I would like to thank Giuseppe Telesca, whom I have cooperated with on several topics, mainly of course on the determination of a sensible Z_{eff} value at TEXTOR. I am very grateful for the many insights he gave me, for instance in appreciating the many sources of uncertainty that affect any measurement, in particular in relation to Z_{eff} . Giuseppe also taught me on impurity transport and the physics of a tokamak plasma in general. I know that he rather preferred for me to pursue the more physics oriented line of research, but yet I hope that he enjoyed as much as I did discussing the many aspects of the bremsstrahlung Z_{eff} diagnostic and the probabilistic estimation of Z_{eff} . And of course I want to thank Giuseppe for the many enjoyable moments in our office in Jülich, where he shared his views with me on the more important things in life as well, such as women.

Ik wil ook graag De Heer H. Schwarz van harte bedanken, die na genoeg de enige directe technische ondersteuning verzorgde ter plaatse in Jülich voor de Gentse spectroscopische groep. Dankzij zijn grote

praktische vaardigheid konden talloze kleine en grotere technische problemen opgelost worden. Reeds vele jaren gepensioneerd, heb ik hem altijd bewonderd om zijn uitstekende conditie, wat vooral tot uiting kwam tijdens de vele keren dat ik mij achter hem aan moest haasten terwijl hij als een volleurde bergbeklimmer door en over TEXTOR klom. Eigenlijk zou ik zijn bedanking in het Duits moeten schrijven, maar zelfs na al die jaren van geregelde bezoeken aan Jülich zou dat nog altijd niet publiceerbaar zijn. Ik ben er evenwel van overtuigd dat De Heer Schwarz dit zal kunnen ontcijferen, uitgaande van zijn theorie dat de Duitse taal langzamerhand overgaat in het Nederlands naarmate men de grens nadert. Het weinige Duits dat ik ken, heb ik van hem geleerd, ondermeer via zijn geregelde levensbeschouwelijke bespiegelingen.

I thank the technical staff of the Institut für Plasmaphysik at Forschungszentrum Jülich for the technical support, especially Mr. Klaus Klören for his help during the calibration of the Z_{eff} -diagnostic and Mr. Jon Krom for his help with the bremsstrahlung Web Umbrella server. Ik wil ook graag Roger Jaspers (FOM-Instituut voor Plasmafysica ‘Rijnhuizen’) bedanken voor het beschikbaar stellen van de vele meters glasvezelkabel die dienen in mijn Z_{eff} -diagnostiek.

The search for a suitable research theme in the field of the determination of Z_{eff} has not been always that easy. I still remember very clearly that it was Manfred von Hellermann who first mentioned to me the enormous potential of various advanced data analysis methods in the field of fusion data processing, such as neural networks, genetic algorithms, etc. This aroused my interest and I am very grateful for his enthusiastic speeches on these subjects, without which I would probably not have started my work on advanced data analysis. I also thank Manfred for all the information he gave me on CXS and the JET CXS system in particular. Ook Kristel Crombé is hier heel behulpzaam in geweest, en samen met Jef Ongena heeft ze ervoor gezorgd dat ik zonder administratieve problemen metingen van JET kon gebruiken; ook daarvoor ben ik dankbaar.

In the initial period of my research in the field of data analysis, I spent a lot of time on the possibilities of neural networks and the related Blind Source Separation methods (BSS, primarily Independent Component Analysis). Although I achieved some fine results using these techniques, still the Bayesian methodology proved to form a better basis. I am very grateful to Dr. Kevin Knuth (NASA Ames) for pointing out the link between BSS and the Bayesian principles. He

introduced me to the workshops on Maximum Entropy and Bayesian methods. I have learned a lot at these workshops, and in turn this put me into contact with Mehdi Ichir from the group of inverse problems at CNRS-Supélec. Mehdi pointed out the possibilities of Markov Chain Monte Carlo and we did some calculations together that indirectly gave rise to part of the work described in this thesis. I am also very grateful to Dr. Rainer Fischer (Max-Planck-Institut für Plasmaphysik) for the further insights he provided me in Bayesian probability theory and computational methods.

Graag wens ik de leden van de examencommissie te danken voor de commentaren en suggesties die ze mij gegeven hebben, en die de kwaliteit van mijn thesis zeker ten goede zijn gekomen.

Ik wil ook een aantal mensen bedanken die geen wetenschappelijke bijdrage geleverd hebben aan mijn werk, maar zonder wie deze thesis en de verschillende jaren van onderzoek niet mogelijk zouden zijn geweest. Dat zijn in de eerste plaats mijn ouders, die mij in alles gesteund hebben en mij dikwijls wijze raad gegeven hebben, terwijl ze steeds voor een aangename huiselijke sfeer zorgden. Mijn vriendin Lien verdient een grote bedanking. Haar lieve bezorgdheid en opgewekte natuur hebben mij heel veel gesteund, en bovendien heeft ze het overgrote deel van de huishoudelijke taken op zich genomen tijdens de verschillende maanden van voorbereiding van mijn thesis. Momenteel beschouw ik de afwas doen als een welkome afwisseling voor Bayesiaanse waarschijnlijkheidsrekening! Ik dank ook de familie van Lien die mij heel goed heeft ontvangen zodat ik mij direct thuis voelde. Daarnaast dank ik mijn zus Annelies en haar man Kalifa voor de steun en aangename momenten, en natuurlijk ook de kleine Zaïna voor de nodige afleiding tijdens het schrijven van mijn thesis. Ten slotte wil ik de rest van mijn familie en mijn vrienden (te beginnen met Kevin, Sarah, Elli en Michael) bedanken voor de vele prettige momenten, waarvan ik hoop dat er nog veel mogen komen.

Geert Verdoolaege
Gent, 17 november 2006

Table of Contents

Dankwoord	i
Table of Contents	v
List of Figures	ix
List of Tables	xvii
List of Acronyms	xix
Nederlandse samenvatting	xxi
1 Het wereldwijde energieprobleem	xxi
2 Gecontroleerde thermonucleaire fusie	xxi
3 Plasmaonzuiverheden	xxiii
4 Experimentele bepaling van Z_{eff}	xxiv
5 Inconsistentie van de continuüm- en CX- Z_{eff}	xxv
6 De zichtbare-remstralingsdiagnostiek aan TEXTOR	xxv
7 Geïntegreerde Data-analyse voor fusiediagnostieken . . .	xxvii
8 Bayesiaanse inferentie	xxvii
9 IDA voor het bepalen van Z_{eff}	xxix
English summary	xxxix
1 The world energy problem	xxxix
2 Controlled thermonuclear fusion	xxxix
3 Plasma impurities	xxxix
4 Experimental determination of Z_{eff}	xxxix
5 Inconsistency of the continuum and CX Z_{eff}	xxxix
6 The visible bremsstrahlung diagnostic on TEXTOR	xxxix
7 Integrated Data Analysis for fusion diagnostics	xxxix
8 Bayesian inference	xxxix

9	IDA for the determination of Z_{eff}	xxxvii
1	Introduction	1
1.1	Topics of interest	1
1.2	Outline	4
1.3	Publications	4
2	Controlled thermonuclear fusion and tokamak physics	9
2.1	The energy problem	9
2.2	Controlled thermonuclear fusion	13
2.3	Tokamaks	16
2.4	Plasma equilibrium	17
2.5	Plasma confinement	18
2.6	Plasma-surface interaction	19
2.7	Limiter and divertor tokamaks	20
2.8	Plasma impurities	24
2.8.1	Harmful and beneficial effects of impurities . . .	25
2.8.2	Wall conditioning	26
2.8.3	Impurity transport	26
2.8.3.1	Edge transport	27
2.8.3.2	Main plasma transport	28
2.8.4	Measuring impurity content	28
2.8.5	The ion effective charge Z_{eff}	29
2.9	Plasma heating	33
2.10	ITER: the next-step tokamak device	34
2.11	Fusion power plants	36
2.12	The TEXTOR tokamak	37
2.13	The JET tokamak	38
3	Spectroscopic determination of Z_{eff}	43
3.1	Visible bremsstrahlung spectroscopy	44
3.1.1	Bremsstrahlung from a plasma	44
3.1.2	Relation with Z_{eff}	47
3.2	CXRS	48
3.2.1	Charge exchange with neutral beam particles . .	48
3.2.2	Calculation of light impurity densities	50
3.2.3	Calculation of Z_{eff}	52
3.2.4	Other parameters derived from CXS	53
3.2.5	A note on data availability	53
3.3	Z_{eff} inconsistency	54

4	Diagnostic systems	61
4.1	Visible bremsstrahlung diagnostic	62
4.1.1	Experimental set-up	62
4.1.1.1	The pre-TEXTOR-DED diagnostic	62
4.1.1.2	The upgraded visible bremsstrahlung diagnostic	63
4.1.2	Calibration	77
4.1.2.1	Integrating spheres	78
4.1.2.2	Calibration procedure	79
4.1.2.3	Calibration difficulties	81
4.1.2.4	Relative calibration estimation from profile consistency	81
4.1.3	Bremsstrahlung emissivity profile reconstruction	89
4.1.4	Z_{eff} calculation	91
4.1.4.1	Line-averaged Z_{eff}	91
4.1.4.2	Z_{eff} profile calculation	92
4.1.4.3	Propagation of edge continuum contributions in Z_{eff} profiles	93
4.2	The CXS diagnostic on JET	96
4.3	Error estimates	96
5	Bayesian probability theory and Bayesian computation	99
5.1	Bayesian probability theory	100
5.1.1	Bayesians versus frequentists	100
5.1.2	Bayes' theorem	104
5.1.3	Marginalization	105
5.1.4	Bayesian model selection and Ockham's razor . .	106
5.1.5	Prior PDF selection	108
5.1.5.1	Uninformative prior distributions . . .	109
5.1.5.2	Conjugate prior distributions	110
5.1.5.3	Encoding expert knowledge	110
5.1.6	Influence of the prior on the posterior	111
5.2	Bayesian computational methods	113
5.2.1	Some historic notes	114
5.2.2	Markov chains	115
5.2.3	The principle of Monte Carlo integration	120
5.2.4	Markov Chain Monte Carlo methods	121
5.2.4.1	The Metropolis-Hastings algorithm . .	122
5.2.4.2	The multivariate Metropolis-Hastings algorithm	126

5.2.4.3	The Gibbs sampler	128
6	Integrated estimation of Z_{eff} from bremsstrahlung and CX spectroscopy	129
6.1	Introduction	129
6.2	Motivation	130
6.3	The BIDA recipe	132
6.4	Z_{eff} IDA: local model	134
6.4.1	Inversion of the local model	138
6.4.2	Inference from the local model: statistical uncertainties	144
6.4.2.1	Model definition	144
6.4.2.2	Posterior calculations	146
6.4.2.3	Consistency analysis	155
6.4.3	Inference from the local model: statistical and systematic uncertainties	159
6.4.3.1	Model definition	159
6.4.3.2	Metropolis sampling	162
6.4.3.3	Consistency analysis	174
6.4.3.4	Conclusion	174
6.5	IDA for the estimation of Z_{eff} profiles	178
6.5.1	Model definition	179
6.5.1.1	Geometry matrix	179
6.5.1.2	Spline representation	181
6.5.1.3	Joint posterior density	182
6.5.1.4	Metropolis sampling	183
7	General conclusions and outlook	189
7.1	Conclusions	190
7.1.1	Visible bremsstrahlung diagnostic	190
7.1.2	Integrated estimation of Z_{eff}	190
7.2	Outlook	193
7.2.1	Visible bremsstrahlung diagnostic	193
7.2.2	Integrated estimation of Z_{eff}	193
	Bibliography	197

List of Figures

2.1	Evolution of the CO ₂ concentration in the atmosphere (in ppm) during the last 1000 years [1].	12
2.2	Annual world energy generation, distributed per energy type (expressed in Million Tons Oil Equivalent), showing also the rise of energy needs during the last decades. . .	13
2.3	Cross-sections of the reactions D-T, D-He ³ and the total cross-section of the two possible D-D reactions.	15
2.4	The progress in fusion research towards reactor conditions, as measured by the product of peak ion density \hat{n} , energy confinement time τ_E and peak ion temperature \hat{T} . . .	16
2.5	The principle of magnetic confinement in a tokamak. . .	17
2.6	Equilibrium flux surfaces and a plot of the toroidal current density j_ϕ , the plasma pressure and the toroidal magnetic field on the midplane.	18
2.7	Several possible limiter and divertor configurations [2]: (i) toroidal limiter, which is analogous to the poloidal divertor in (ii), the poloidal limiter (iii) and toroidal divertor (iv), the rail limiter (v) and bundle divertor (vi).	22
2.8	Schematic showing the minor radius a and major radius R_0 for a tokamak with a circular poloidal cross-section.	23
2.9	Schematics of a divertor configuration, where the poloidal field \mathbf{B}_θ is diverted by the divertor coil, internal or external to the vacuum vessel.	23
2.10	The divertor SOL, showing also the divertor target plates.	24
2.11	The definition of the plasma edge in the context of transport, with indicated ionization depth $\lambda_{i,Z}$	27
2.12	Sketch of the various heating methods for a tokamak plasma.	34
2.13	ITER cutaway, showing the major components of the tokamak itself.	35
2.14	A schematic of a tokamak fusion power station.	36

2.15	Inside view of TEXTOR, with parts of the ALT-II limiter clearly visible on the lower right-hand side, and the graphite tiles covering the DED coils on the left-hand side.	39
2.16	Poloidal cross-section of a typical lay-out of the JET magnetic flux surfaces.	40
2.17	Inside view of the JET vacuum vessel, with the divertor chamber clearly visible at the bottom.	40
3.1	The free-free Gaunt factor \bar{g}_{ff} calculated with the approximations of Griem, Elwert, Kramer and Ramsey for $Z_i = 6$ at $\lambda = 5230 \text{ \AA}$. The result of the numerical calculations by Karzas <i>et al.</i> is also shown.	47
3.2	Schematic of a possible geometry of lines of sight observing a neutral beam.	49
3.3	Charge exchange spectrum of CVI at 5290.5 \AA ($T_e = 19.3 \text{ keV}$) from the horizontal JET CX diagnostic (see Section 4.2). Note the active (beam-related) and passive contributions to the spectrum ($T_e = 3.6 \text{ keV}$) [3].	51
3.4	Line-averaged continuum Z_{eff} and equivalent CX Z_{eff} line integral for JET #42982.	55
3.5	On-axis continuum Z_{eff} and CX Z_{eff} in a carbon-dominated plasma for JET #60718.	56
3.6	On-axis continuum Z_{eff} and CX Z_{eff} in a carbon-dominated plasma for JET #61352.	56
4.1	Overview of the main components of the visible bremsstrahlung Z_{eff} diagnostic on TEXTOR.	64
4.2	Impact factor and intercept (relative to the the centre of the vacuum vessel) for an example chord viewing a plasma cross-section.	65
4.3	Top view of the TEXTOR vessel, with the port for the visible bremsstrahlung Z_{eff} diagnostic indicated (' Z_{eff} ').	66
4.4	The viewing geometry of the TEXTOR visible bremsstrahlung diagnostic in a poloidal cross-section. The red lines indicate the field of view containing the 31 lines of sight.	67
4.5	Part of the TEXTOR line spectrum, indicating no spectral lines in the immediate vicinity of λ_{ff} . The wavelengths are indicated in \AA	68

4.6	The principle of imaging and charge read-out in a CCD sensor, with the horizontal register at the bottom of the array.	70
4.7	A screenshot of the interface to the control and image processing software ZeffControl of the visible bremsstrahlung Z_{eff} diagnostic on TEXTOR.	72
4.8	Removing the smearing effect in the CCD images.	77
4.9	The integrating sphere mounted inside the TEXTOR vessel. The viewing geometry of the lines of sight has been indicated on a strip of paper.	80
4.10	Influence of an erroneous calibration on the line-integrated emissivity profile. The data points are interconnected to facilitate recognizing the similarities in both curves. . .	82
4.11	Simulated calibrated line-integrated emissivity profiles, original (blue line) and after the plasma shift has occurred (green line). The calibration factor of channel 10 has artificially been altered from the true one.	85
4.12	The geometry of a chord during a horizontal plasma shift.	87
4.13	A simulation of the relative calibration method based on (4.3). The blue curve represents the original artificial line-integrated bremsstrahlung emissivity profile, the red curve is the profile based on the calculated calibration. The profile based on the calculated calibration was rescaled and shifted back to match the original profile.	88
4.14	The calibrated line-integrated bremsstrahlung profiles resulting from the requirement of profile consistency under a horizontal plasma shift H . H is indicated in the figure for two channels.	88
4.15	Reconstructed Z_{eff} profiles using Abel inversion, Tikhonov and Maximum Entropy regularized inversion for TEXTOR discharge #99434 at 1.9 s. The Z_{eff} overestimation at the plasma edge is the largest for the Abel inverted profile.	92
4.16	Time traces of respectively: electron density, DED current, CIII, CV and CVI line intensity, line-averaged Z_{eff} and edge and central Z_{eff} for TEXTOR discharge #99433. From the carbon line intensity signals, a decontamination or screening effect due to DED can be noticed. However, the central Z_{eff} appears to be unaffected by DED.	94
4.17	Influence of DED on the Z_{eff} profile for TEXTOR discharge #99433.	95

5.1	Illustration of the quantities that determine the Ockham factor for a model \mathcal{H}_i with a single parameter θ . The prior distribution for the parameter is uniform (dotted line) and has a width $\Delta^0\theta$. The posterior (solid line) has a single maximum at θ_{MP} with characteristic width $\Delta\theta$.	108
5.2	Histograms (after burn-in) for three random walk Metropolis simulations on $[-\delta, \delta]$ with (a) $\delta = 0.1$, (b) $\delta = 0.5$ and (c) $\delta = 1.0$. The Monte Carlo estimates of the means throughout the simulation have been superimposed (scale on the right side of the graphs).	127
6.1	Time traces for the on-axis n_e , respectively T_e , ϵ_{ff} , n_C , ϵ and δ for JET pulse #60718. The units for ϵ_{ff} have been abbreviated, and are $\text{W} / (\text{cm}^2 \text{ sr } \text{\AA})$.	139
6.2	Time traces for the on-axis n_e , respectively T_e , ϵ_{ff} , n_C , ϵ and δ for JET pulse #61352. The units for ϵ_{ff} have been abbreviated, and are $\text{W} / (\text{cm}^2 \text{ sr } \text{\AA})$.	140
6.3	Time traces for the on-axis density from LIDAR and the calculated density through model inversion for pulse #60718.	141
6.4	Time traces for the on-axis continuum Z_{eff} and CX Z_{eff} , and the calculated Z_{eff} through model inversion for pulse #60718.	142
6.5	Time traces for the on-axis density from LIDAR and the calculated density through model inversion for pulse #61352.	142
6.6	Time traces for the on-axis continuum Z_{eff} and CX Z_{eff} , and the calculated Z_{eff} through model inversion for pulse #61352.	143
6.7	Contour plot for the joint posterior density of the on-axis n_e and Z_{eff} for JET #61352 at 47s. The numbers indicate probability density.	147
6.8	The joint posterior density of the on-axis n_e and Z_{eff} for JET #61352 at 47s. The numbers indicate probability density.	148
6.9	Marginal posterior distributions for the on-axis n_e and Z_{eff} for JET #61352 at 47s. Note how the probability density is cut off at $Z_{\text{eff}} = 1$.	149

6.10	Marginal posterior means of the on-axis n_e for JET #61352. The error bars on the posterior means correspond to a single standard deviation.	150
6.11	Marginal posterior means of the on-axis Z_{eff} for JET #61352. The error bars on the posterior means correspond to a single standard deviation.	150
6.12	Marginal posterior distributions for the on-axis n_e and Z_{eff} for JET #61352 at 47s. The density from LIDAR is taken as an additional measurement.	151
6.13	Marginal posterior means of the on-axis n_e for JET #61352, including also LIDAR measurements.	152
6.14	Marginal posterior means of the on-axis Z_{eff} for JET #61352, including also LIDAR measurements.	152
6.15	Marginal posterior means of the on-axis n_e for JET #60718, including also LIDAR measurements.	153
6.16	Marginal posterior means of the on-axis Z_{eff} for JET #60718, including also LIDAR measurements.	154
6.17	Marginal posterior means of the on-axis Z_{eff} for JET #61352, including also LIDAR measurements with an increased error of 20%.	156
6.18	Marginal posterior means of the on-axis Z_{eff} for JET #61352, including also LIDAR measurements with an increased error of 20%. The error on δ has also been increased to 50%.	156
6.19	Marginal posterior means of the on-axis Z_{eff} for JET #60718, including also LIDAR measurements with an increased error of 20%. The error on δ has also been increased to 50%.	157
6.20	Marginal posterior distributions for the on-axis n_e for JET pulse #61352 at 47s, taking into account different data sets.	159
6.21	Zoomed display of Figure 6.21.	160
6.22	Marginal posterior distributions for the on-axis Z_{eff} for JET pulse #61352 at 47s, taking into account different data sets.	160
6.23	The marginal posterior for the on-axis n_e for the combination of all three likelihoods.	161
6.24	The marginal posterior for the on-axis Z_{eff} for the combination of all three likelihoods.	161
6.25	The standard Cauchy distribution $C(\theta 0, 1)$	164

6.26	Histogram of the sampled marginal distribution of the on-axis Z_{eff} after a Monte Carlo run of 30,000 samples, using an artificially imposed uniform posterior distribution ($1 \leq Z_{\text{eff}} \leq 5$).	165
6.27	Monte Carlo mean of the on-axis n_e using an artificial data set.	166
6.28	Monte Carlo mean of the on-axis Z_{eff} using an artificial data set.	166
6.29	Autocorrelation between n_e samples in an MCMC chain as a function of time lag s	168
6.30	MCMC trace of the sampled on-axis n_e and Z_{eff} at 63 s in JET #60718. There is a clear effective burn-in of about 1000 samples.	169
6.31	Joint posterior probability density in an MCMC run using data from JET #60718.	169
6.32	Monte Carlo mean of the on-axis n_e using data from JET #60718.	170
6.33	Monte Carlo mean of the on-axis Z_{eff} using data from JET #60718	170
6.34	Monte Carlo mean of the on-axis n_e using data from JET #61352.	171
6.35	Monte Carlo mean of the on-axis Z_{eff} using data from JET #61352.	172
6.36	Monte Carlo mean of the on-axis n_e using data from JET #59193.	172
6.37	Monte Carlo mean of the on-axis Z_{eff} using data from JET #59193.	173
6.38	Monte Carlo mean of the on-axis n_e using data from JET #59186.	173
6.39	Monte Carlo mean of the on-axis Z_{eff} using data from JET #59186.	174
6.40	Monte Carlo mean of the on-axis n_e using data from JET #59194.	175
6.41	Monte Carlo mean of the on-axis Z_{eff} using data from JET #59194.	175
6.42	Marginal posterior distributions for the on-axis n_e for JET pulse #61352 at 47s, taking into account different data sets. The ϵ and δ measurements have been scale corrected.	176

6.43	Marginal posterior distributions for the on-axis Z_{eff} for JET pulse #61352 at 47s, taking into account different data sets. The ϵ and δ measurements have been scale corrected.	176
6.44	Schematic representation of a simple circularly-symmetric magnetic configuration (the Shafranov shift has been neglected), traversed by a horizontal line of sight.	181
6.45	An artificial electron density profile used for the MCMC estimation of n_e and Z_{eff} profiles.	184
6.46	An artificial electron temperature profile used for the MCMC estimation of n_e and Z_{eff} profiles.	184
6.47	An artificial Z_{eff} profile used for the MCMC estimation of n_e and Z_{eff} profiles.	185
6.48	The eight spline basis functions that were used in the MCMC estimation of n_e and Z_{eff} profiles.	185
6.49	MCMC electron density profile estimate from an artificial data set consisting of line-integrated bremsstrahlung emissivity, local CX impurity density and line-integrated electron density. The profile was approximated by a spline.	186
6.50	MCMC Z_{eff} estimate from an artificial data set consisting of line-integrated bremsstrahlung emissivity, local CX impurity density and line-integrated electron density. The profile was approximated by a spline.	187

List of Tables

2.1	Average annual per capita total primary power consumption for selected countries (figures date from 1995) [4]. . .	10
2.2	Estimated maximum years of use of different fuels at the current rate of consumption [4].	11
2.3	Investments needed to produce 1 GW of electricity for several types of renewables [4]. Surfaces can be compared to the total surface of Belgium: 32.545 km ²	14
2.4	Typical values of some of the principal parameters for TEXTOR, JET and ITER.	41
4.1	The main operational parameters of the CCD camera used for the visible bremsstrahlung diagnostic on TEXTOR.	71
4.2	The estimated relative errors on several local plasma quantities in JET or TEXTOR plasmas. The errors are generally both of a statistical and systematic nature. . .	95
5.1	Monte Carlo estimates of the mean and variance of a normal distribution $N(0, 1)$ based on a sample (after burn-in) from a Metropolis chain using a random walk on $[-\delta, \delta]$	126
6.1	Scientific programme associated with the JET discharges under study in this work.	137
6.2	MCMC estimated scale factors for the on-axis ϵ and δ in several JET discharges.	178

List of Acronyms

ADC	Analog-to-Digital Convertor
ADU	Analog-to-Digital Units
ALT	Advanced Limiter Test
BIDA	Bayesian Integrated Data Analysis
BSS	Blind Source Separation
CCD	Charge-Coupled Device
CFC	Carbon Fibre Composite(s)
CSF	Common Storage Facility
CX	Charge Exchange
CXRS	Charge Exchange Recombination Spectroscopy
CXS	Charge Exchange Spectroscopy
DED	Dynamic Ergodic Divertor
ECRH	Electron Cyclotron Resonance Heating
FWHM	Full Width at Half Maximum
HFS	High Field Side
ICA	Independent Component Analysis
ICRH	Ion Cyclotron Resonance Heating
IDA	Integrated Data Analysis
ITER	International Thermonuclear Experimental Reactor
JET	Joint European Torus

LCFS	Last Closed Flux Surface
LFS	Low Field Side
LHRH	Lower Hybrid Resonance Heating
LIDAR	Light Detection and Ranging
MAP	Maximum A Posteriori
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
M-H	Metropolis-Hastings
MHD	magnetohydrodynamic(s)
NBI	Neutral Beam Injection
PDF	Probability Density Function
PM	photomultiplier
SOL	Scrape-Off Layer
TEC	Trilateral Euregio Cluster
TEXTOR	Torus Experiment for Technology Oriented Research
TPD	TEXTOR Physics Database
TTL	Transistor-Transistor Logic
TWU	TEC Web Umbrella

Nederlandse samenvatting —Summary in Dutch—

1 Het wereldwijde energieprobleem

Om de levensstandaard van de gemiddelde inwoner van een modern geïndustrialiseerd land te onderhouden, is een aanzienlijke hoeveelheid energie nodig. Vandaag wordt de meeste energie opgewekt door de verbranding van fossiele brandstoffen en door kernsplijting. De belangrijkste reserves aan fossiele brandstoffen en uranium zullen echter uitgeput zijn binnen de 50 jaar (voor petroleum en uranium) tot 250 jaar (voor steenkool). Bovendien veroorzaakt het verbranden van fossiele brandstoffen ernstige milieuproblemen, onder andere als gevolg van de uitstoot van broeikasgassen. Er lijkt nu een algemene consensus te bestaan dat dit globale klimaatwijzigingen veroorzaakt, veranderingen die zelfs onomkeerbaar zouden kunnen zijn. Anderzijds is het veilig hanteren en opslaan van kernaafval, afkomstig van kernsplijting, eveneens een complexe aangelegenheid.

Hoewel er verschillende hernieuwbare energiebronnen bestaan, zoals zonne- en windenergie, eisen deze methodes gemiddeld enorm veel ruimte op, en het is nu reeds duidelijk dat deze energiebronnen enkel een aanvulling kunnen vormen op bestaande en toekomstige schone energiebronnen.

2 Gecontroleerde thermonucleaire fusie

Een mogelijke oplossing voor dit probleem wordt geboden door gecontroleerde thermonucleaire fusie². Kernfusie is het proces dat de sterren doet stralen. Het verwijst naar het samensmelten van atoomkernen, waarbij het massadefect in equivalente energie vrijkomt. Het

²Voor een populariserende inleiding tot gecontroleerde kernfusie verwijzen we naar Ref. [5].

onderzoek naar gecontroleerde fusie op aarde heeft tot doel energie te produceren op basis van kernfusie. Om dit te bereiken wordt de fusie-reactie tussen een deuterium- (D) en een tritiumkern (T) beschouwd, wat resulteert in een heliumkern en een neutron; het is dit neutron dat de meeste fusie-energie met zich mee draagt. Als gevolg van de Coulomb-barrière tussen twee kernen vraagt deze reactie erg veel energie. Rekening houdend met kwantummechanische tunneling is dit in de orde van 10 keV, het equivalent van 100 miljoen graden Celsius. Bij zulke energieën worden alle D- en T-kernen geïoniseerd. Het geheel van brandstofionen en elektronen wordt een *plasma* genoemd, wat de vierde aggregatietoestand is.

Het opsluiten van (experimentele) fusieplasma's kan niet gebeuren door een materiële wand. In de plaats daarvan wordt het plasma gevangen in een magnetisch veld. De verst ontwikkelde configuratie op dat vlak is de torusvormige *tokamak*. Het magnetisch veld in een tokamak is een superpositie van een toroïdaal veld, aangelegd door toroïdale spoelen, en een poloïdaal veld gegenereerd door een grote inductieve plasmastroom (honderden kA tot zelfs MA). De geladen plasmadeeltjes gyreren rond de veldlijnen, en zijn op die manier, tot op zekere hoogte, opgesloten in het vacuümvat van de tokamak. Desondanks is er ook een radiaal transport van deeltjes en energie als gevolg van diffusie-, convectie- en stralingsprocessen, wat aanleiding geeft tot een eindige deeltjes- en energieopsluitingstijd. Om thermonucleaire condities te bereiken, moeten de dichtheid van de brandstof, de temperatuur en de opsluitingstijd tegelijkertijd voldoende hoog zijn. Het plasma wordt verhit door middel van verschillende technieken: Joule-verhitting door de plasmastroom, evenals aanvullende verhitting door de injectie van energierijke waterstofachtige atomen (Neutrale-Bundelinjectie), en door de resonante absorptie van radiogolven. Uiteindelijk moet het vermogensverlies in het plasma gecompenseerd worden door de verhitting vanwege de alfadeeltjes. Op dat moment is geen aanvullende plasma-verhitting meer vereist. Dit gebeuren wordt *ontbranding* genoemd.

Het huidige onderzoek werd gedeeltelijk uitgevoerd aan de TEXTOR-tokamak (Institut für Plasmaphysik, Forschungszentrum Jülich, Duitsland) en aan JET (EFDA, Culham, UK), de grootste tokamak ter wereld. Een aanzienlijk deel van de wereldwijde inspanningen rond tokamakonderzoek is geconcentreerd op de toekomstige ITER-machine. Het doel van ITER is om de wetenschappelijke en technische haalbaarheid aan te tonen van energie uit kernfusie voor vredelievende doeleinden. De start van de eigenlijke constructie van ITER is gepland in de loop van

2008, terwijl het eerste ITER-plasma verwacht wordt in 2016. ITER zal de enige stap zijn tussen de huidige tokamaks en de eerste fusiereactor DEMO.

3 Plasmaonzuiverheden

Hoewel er de afgelopen decennia een aanzienlijke vooruitgang is geboekt in het onderzoek naar gecontroleerde kernfusie, zijn er nog steeds een aantal kwesties die de realisatie van de thermonucleaire condities belemmeren. Eén van de belangrijkste moeilijkheden heeft te maken met de controle van de ionen in het plasma die niet tot de brandstof behoren, de zogenaamde *onzuiverheden*. Onzuiverheden worden losgelaten van componenten die in contact staan met het plasma, en dit door verschillende plasma-wand-interactieprocessen, zoals sputtering. Geïoniseerde onzuiverheden kunnen propageren naar het centrum van het plasma. Daar geven ze aanleiding tot verdunning van de brandstof, terwijl ze de energieopsluiting verminderen door middel van remstraling. Dit zijn de belangrijkste nadelige eigenschappen van onzuiverheden, die, bij reactorvoorwaarden, de ontbranding kunnen verhinderen. Anderzijds kan in de nabijheid van de rand van het plasma atomaire lijnstraling door onzuiverheden de plasma-wand-interactie verlagen, zonder afbreuk te doen aan de energieopsluiting in het centrum van het plasma. Met dat oogmerk kunnen lichte onzuiverheden zelfs geïnjecteerd worden. Samen met de mogelijkheden van onzuiverheden voor plasmadiagnose is dit één van de belangrijkste positieve effecten van onzuiverheden.

Om het gedrag te karakteriseren en de gevolgen af te leiden van onzuiverheden, is het cruciaal om informatie te bekomen over het onzuiverheidsgehalte van het plasma. Dit kan gekwantificeerd worden door de zogenaamde *effectieve ionaire lading* Z_{eff} , de grootte die een sleutelrol speelt in dit werk. Z_{eff} wordt gedefinieerd als

$$Z_{\text{eff}} = \frac{\sum_i n_i Z_i^2}{\sum_i n_i Z_i}, \quad (1)$$

waarbij n_i de dichtheid is voor de onzuiverheidssoort (of waterstofachtige soort) i , en Z_i zijn lading. Z_{eff} is een lokale maat voor de onzuiverheidsconcentratie, gemiddeld over alle onzuiverheden. Z_{eff} kan rechtstreeks in verband gebracht worden met de elektrische resistiviteit van het plasma. Voor ITER wordt een Z_{eff} -waarde van 1,8 voorzien. Af-

hankelijk van het ontladingsscenario mag dit getal slechts variëren met $\pm 0, 2$.

4 Experimentele bepaling van Z_{eff}

Er bestaan verschillende methodes voor het afleiden van Z_{eff} . In het huidige werk wordt Z_{eff} berekend op basis van remstralingsspectroscopie aan de ene kant, en ladingsuitwisselingspectroscopie aan de andere kant. De remstraling in een plasma vindt haar oorsprong vooral in de vrije overgang van een elektron in het elektrisch veld van een ion. Eénmaal de elektronendichtheid n_e en elektronentemperatuur T_e bekend zijn, is de lokale remstralingsemissiviteit ϵ_{ff} (gewoonlijk in $\text{W}/\text{cm}^3/\text{sr}/\text{\AA}$) recht evenredig met Z_{eff} :

$$\epsilon_{\text{ff}} \sim \bar{g}_{\text{ff}}(T_e) \frac{n_e^2 Z_{\text{eff}}}{\sqrt{T_e}},$$

waarbij \bar{g}_{ff} de Maxwell-gemiddelde Gaunt-factor is, die alle kwantummechanische effecten omvat. Een remstralingsdiagnostiek observeert lijnintegralen van remstralingsemissiviteit, en een radiaal profiel voor de lokale emissiviteit moet afgeleid worden door middel van een inversieprocedure. Indien de profielen van n_e en T_e eveneens beschikbaar zijn, kan vervolgens een Z_{eff} -profiel berekend worden.

Bij de typische energieën die heersen in het centrum van tokamakplasma's, zijn de meeste ionen volledig geïoniseerd. Dat maakt de directe spectroscopie van onzuiverheden onmogelijk. De energetische waterstofachtige atomen van een neutrale bundel kunnen evenwel een ladingsuitwisselingsreactie veroorzaken met een plasma-ion. Het plasma-ion ontvangt een elektron in een geëxciteerde toestand, terwijl het geïoniseerde bundeldeelje de bundel verlaat. Door het observeren van het stralingsverval van de onzuiverheidsionen in het plasma, is het vervolgens mogelijk de dichtheid van de onzuiverheden af te leiden, samen met de ionentemperatuur en plasmastroming. Dit is het doel van *ladingsuitwisselingspectroscopie*³. Indien de belangrijkste plasma-onzuiverheden gevolgd worden, kan Z_{eff} afgeleid worden op basis van (1). Zowel op TEXTOR als op JET is koolstof de dominante onzuiverheid, zodat vaak alleen de dichtheid van volledig geïoniseerd koolstof in overweging genomen wordt in de sommatie (1). Dit is ook de benadering die genomen wordt in het huidige werk. Elke zichtlijn van

³Charge Exchange Spectroscopy, of CXS

een CXS-diagnostiek observeert een vrij goed gelokaliseerd plasmavolume, gedefinieerd door de bundelgeometrie. Op die manier is geen inversieprocedure meer nodig om de Z_{eff} -profielen af te leiden.

5 Inconsistentie van de continuüm- en CX- Z_{eff}

Vele bronnen van onzekerheid komen voor bij het afleiden van Z_{eff} , zowel uit remstralingsspectroscopie (continuüm- Z_{eff}) als uit CXS (CX- Z_{eff}). Deze onzekerheden kunnen van statistische of systematische aard zijn, en ze kunnen leiden tot aanzienlijke fouten op de berekende Z_{eff} -waarden. Inderdaad, in het algemeen is de continuüm- Z_{eff} inconsistent met de CX- Z_{eff} . Dit is een oud probleem in fusiediagnose en de reconstructie van een Z_{eff} -profiel dat betrouwbaar is over de volledige plasmadoorsnede, is momenteel een echte uitdaging. Eén van de doelstellingen van dit werk is het schatten van een Z_{eff} -waarde die consistent is met zowel de remstralingsemissiviteit als de CX-metingen van onzuiverheidsdichtheid.

6 De zichtbare-remstralingsdiagnostiek aan TEXTOR

Een tweede component van dit werk heeft betrekking op het ontwerp en de ingebruikname van een nieuwe diagnostiek voor remstralings-spectroscopie in het zichtbare aan TEXTOR. Het hoofdsysteem bestaat uit een set van 24 zichtlijnen die een deel van een poloïdale doorsnede van het plasma bekijken. Een antireflectielaag vermindert reflecties binnenin het vat. Het licht wordt door middel van glasvezeloptica getransporteerd naar een diagnostische ruimte voorbij het biologische schild. Een Fabry-Pérot-interferentiefilter selecteert een golflengteband die vrij is van lijnstraling op TEXTOR. Het licht wordt gefocuseerd op een gekoelde CCD-matrix. De CCD wordt normaal op 0°C gehouden, waar de typische signaal-ruis-verhouding voor een belichtingstijd van 50 ms en een pixelgroepering van 8×8 , tegen een uitleessnelheid van 4.0 Mpixel/s, gelijk is aan 45 dB. Gedurende een typische data-acquisitie-sequentie wordt een serie beelden genomen met een herhalingsfrequentie van 20 Hz. Op die manier slaagt de diagnostiek erin om een verhoogde tijdsresolutie te realiseren, vergeleken met het voorgaande systeem dat geïnstalleerd was aan TEXTOR, waarbij er tegelijkertijd een groot aantal koorden is. De tijdregeling van de diagnostiek

wordt gecontroleerd door een programmeerbaar tijdcircuit. De beelden worden verstuurd naar een werkstation voor verdere verwerking, wat resulteert in een set van lijngeïntegreerde tijdssignalen van remstraling. In het kader van dit onderzoek werd een aangepaste software-interface geschreven en een programma om de diagnostiek te controleren en de data te verwerken. Het uitsmeren van CCD-beelden wordt geëlimineerd door de bijdrage door uitsmering te schatten in een donker gedeelte van het beeld, zonder de noodzaak de puntspreidingsfunctie van de optische opstelling te berekenen. Een set glasvezelkabels uitgerust met fotomultiplicatorbuizen werd geïnstalleerd als reserve voor het camerasysteem.

De data worden doorgestuurd naar de centrale opslagfaciliteit van TEXTOR. Een lijngemiddeld tijdsignaal voor Z_{eff} wordt routinematig berekend voor een centraal observerende zichtlijn, en is on-line beschikbaar via de TEXTOR Physics Database. Z_{eff} -profielen kunnen worden gereconstrueerd door een Abel-inversie. Dit geeft vaak aanleiding tot een divergent Z_{eff} -profiel nabij de plasmarand, wat toegeschreven kan worden aan niet-remstralingsbijdragen tot de continuümstraling. Een aantal tests zijn uitgevoerd met Tikhonov- en Maximum-Entropie-geregulariseerde inversie, resulterend in een aanzienlijke reductie van de overschatting van Z_{eff} aan de rand.

De kalibratie van het systeem werd uitgevoerd door gebruik te maken van een integrerende sfeer, die een uniforme en constante lichtbron verschaft. Tot nu toe leidde deze procedure echter nog niet tot een bevredigende relatieve (kanaal-tot-kanaal) kalibratie. Daarom werd een alternatieve techniek ontwikkeld om de relatieve kalibratie te bepalen, gebaseerd op de vereiste van consistentie van lijngeïntegreerde remstralingsprofielen onder een horizontale verschuiving van de plasmakolom. Deze kalibratie leidt tot fysisch aanvaardbare Z_{eff} -profielen. Bovendien is deze methode algemeen toepasbaar: ze voorziet in een eenvoudige en zelfconsistente methode voor de relatieve kalibratie van ook andere meerkanaals spectroscopische diagnostieken. Er is geen specifieke lichtbron of een andere gespecialiseerde kalibratieopstelling nodig, noch is toegang tot vacuümvat vereist. Vooral dit laatste is belangrijk omdat frequente in-situ-kalibraties steeds moeilijker zullen worden in toekomstige fusiemachines.

7 Geïntegreerde Data-analyse voor fusiediagnostieken

De doelstelling van plasmadiagnose is het afleiden van betrouwbare schattingen voor de fysische grootheden die van interesse zijn, vertrekkend van de ruwe opgemeten data. Een kernfusie-experiment genereert grote hoeveelheden ruwe data die op een efficiënte manier verwerkt moeten worden. Dit is de fase van datavalidatie en -analyse. Dit proces houdt mogelijk de integratie in van data van heterogene oorsprong. De afgeleide fysische grootheden worden aangetast door verscheidene bronnen van onzekerheid, en daardoor is een probabilistische analyse aangewezen. Dit leidt naar het domein van de Geïntegreerde Data-Analyse⁴, waar metingen van verschillende diagnostieken gecombineerd kunnen worden, eventueel samen met informatie over het onderliggend fysisch model. In vergelijking met de inspanningen die geleverd worden voor het verfijnen van de methodes voor plasmadiagnose en het verbeteren van diagnostische apparatuur, is IDA vaak een nevenactiviteit. Nochtans kan deze consistente verwerking van alle beschikbare data de betrouwbaarheid en de robuustheid van fysische resultaten gevoelig verbeteren. Bovendien kan IDA aanzienlijk bijdragen tot de optimalisering van het design van diagnostieken. Daarenboven zal de beschikbare ruimte voor diagnostische opstellingen aan ITER beperkt zijn, waardoor elk type van beschikbare informatie uitgebuit zal moeten worden voor de bepaling van fysische grootheden.

8 Bayesiaanse inferentie

In dit werk wordt IDA gevoerd binnen een Bayesiaans probabilistisch kader (vandaar *Bayesiaanse Geïntegreerde Data-Analyse*⁵). Het Bayesiaans paradigma beschouwt een probabilliteit als een graad van waarschijnlijkheid, niet als een frequentie. Bayesiaanse waarschijnlijkheidsrekening is uitstekend geschikt voor het afleiden van de onderliggende parameters van een fysisch model, gegeven een dataset, mogelijk aangevuld door a priori-expertise met betrekking tot de betreffende parameters. De grondslag van Bayesiaanse inferentie is het theorema van Bayes, dat een mogelijkheid biedt om voorwaardelijke probabiliteiten

⁴Integrated Data Analysis, of IDA

⁵Bayesian Integrated Data Analysis, kortweg BIDA

om te keren:

$$\underbrace{p(\boldsymbol{\theta}|\boldsymbol{x}, I)}_{\text{a posteriori}} = \frac{\overbrace{p(\boldsymbol{x}|\boldsymbol{\theta}, I)}^{\text{likelihood}} \overbrace{p(\boldsymbol{\theta}|I)}^{\text{a priori}}}{\underbrace{p(\boldsymbol{x}|I)}_{\text{bewijs}}},$$

waar \boldsymbol{x} een datavector voorstelt, $\boldsymbol{\theta}$ zijn de parameters van interesse, en I beduidt elke vorm van bijkomende beschikbare informatie.

Om in een meerparametermodel gevolgtrekkingen te maken met betrekking tot individuele parameters, moet de gezamenlijke a posteriori-distributie gemarginaliseerd worden. Dit is eveneens vereist voor de berekening van a posteriori momenten. De multidimensionale integralen kunnen moeilijk op te lossen zijn, zelfs numeriek, waardoor men een beroep moet doen op stochastische integratiemethodes. In het huidige werk wordt een Markov-keten-Monte Carlo-schema⁶ gebruikt, meer bepaald het Metropolis-Hastings-steekproefalgoritme. Dit omvat het opzetten van een Markov-keten die convergeert naar de doel-distributie. De schattingen van grootheden van interesse worden afgeleid door het berekenen van het Monte Carlo-gemiddelde (ergodisch gemiddelde) van de bemonsterde parameters.

Het BIDA-recept kan nu als volgt uiteengezet worden. Eerst moeten de belangrijkste bronnen van onzekerheid in het data-beschrijvende model geïdentificeerd worden. Statistische onzekerheden worden gemodelleerd in de variantie van gepaste zogeheten ‘likelihood-distributies’, terwijl systematische onzekerheden voorgesteld worden door zogenaamde *storende parameters*. Dit zijn parameters die niet van interesse zijn, maar die noodzakelijkerwijze in het rechtstreekse model⁷ voorkomen. Vervolgens wordt a priori-informatie gekwantificeerd en een gezamenlijke distributie wordt geformuleerd op basis van het theorema van Bayes. Storende parameters worden uit-geïntegreerd, en marginale momenten voor de parameters van interesse kunnen worden berekend door gebruik te maken van, bijvoorbeeld, MCMC. Uiteindelijk maakt een gevoeligheidsanalyse het mogelijk om de impact te bestuderen van de verschillende bronnen van onzekerheid op de foutenmarges van de grootheden van interesse. Samen met technieken van Bayesiaanse Diagnostische Design, kan het ontwerp van de diagnostieken geoptimaliseerd worden.

⁶Markov Chain Monte Carlo, of MCMC

⁷forward model

9 IDA voor het bepalen van Z_{eff}

In het huidige werk wordt IDA gebruikt voor de schatting van een Z_{eff} -waarde die consistent is met zowel de metingen van remstralingsemissiviteit als met de onzuiverheidsdichtheid uit CXS. Dit wordt toegepast op data van JET-ontladingen, aangezien de CXS-database aan JET veel uitgebreider is dan die aan TEXTOR. Bovendien was de CXS-diagnostiek aan TEXTOR niet recent gekalibreerd toen dit onderzoek uitgevoerd werd. De remstralingsemissiviteit wordt bepaald uit de continuüm-achtergrond van het CX-spectrum. In een eerste fase wordt een Bayesiaans probabilistisch model geformuleerd voor het schatten van (tijdssignalen van) Z_{eff} op de magnetische as. De input-data zijn de lokale remstralingsemissiviteit (en de elektronentemperatuur) aan de ene kant, en de lokale koolstofdichtheid verkregen uit CXS aan de andere kant. Dit model maakt het ook mogelijk om de elektronendichtheid n_e te schatten, die vervolgens vergeleken kan worden met de dichtheid uit LIDAR-Thomson-verstrooiing. Dit geeft een mogelijkheid om de geldigheid van de berekeningen te controleren.

Vooreerst worden enkel de statistische onzekerheden op de data gemodelleerd in termen van Gaussiaanse likelihood-distributies, terwijl niet-informatieve uniforme a priori-distributies verondersteld worden. De analyse wordt toegepast op data van twee JET-ontladingen, éénmaal zonder en éénmaal met de expliciete insluiting van LIDAR-metingen. Het model is voldoende eenvoudig om de berekening toe te laten van de gezamenlijke a posteriori-distributie op een rooster van parameterwaarden. Een consistente Z_{eff} wordt verkregen, maar de geldigheid van de resultaten wordt in twijfel getrokken in gevallen waar er een grote systematische discrepantie bestaat tussen de continuüm- en de CX- Z_{eff} . Dit wordt bevestigd in een consistentieanalyse.

Daarom wordt een tweede model voorgesteld, waar ook de systematische onzekerheden gemodelleerd worden, door middel van een schaalfactor voor elk van de metingen. De a priori-distributie voor de schaalfactoren is een uniforme distributie. Schattingen voor de parameters van interesse worden verkregen via MCMC, namelijk een zuiver Metropolis-steekproef-schema met een Cauchy-voorgestelde distributie⁸. De Cauchy-schalen worden zó afgesteld dat een voordelige aanvaardings-frequentie⁹ verkregen wordt. De methode wordt eerst getest op een artificiële dataset, en vervolgens op data van een aantal

⁸proposal distribution

⁹acceptance rate

JET-ontladingen. Een consistentieanalyse toont aan dat de geschaalde input-data onderling veel consistentier zijn dan hun niet-geschaalde versies. Het steekproefalgoritme functioneert goed en consistente Z_{eff} -schattingen kunnen worden bepaald, met aanzienlijk gereduceerde foutenmarges vergeleken met de fouten op de continuüm- en CX- Z_{eff} . In verschillende van de ontladingen die bestudeerd worden, is het geschatte Z_{eff} -tijdsignaal gelijkaardig, zowel kwalitatief als kwantitatief, met de CX- Z_{eff} , eerder dan met de continuüm Z_{eff} . De schaalfactor voor de continuümmeting ligt systematisch boven de eenheid, wat wijst op een overschatting van de remstralingsemissiviteit met een gemiddelde factor van 1,4. In enkele van de beschouwde ontladingen wordt de CX-meting, d.i. de gesommeerde onzuiverheidsdichtheid, onderschat met een factor van 0,4 tot 0,8. Deze resultaten suggereren het opzetten van een databank van MCMC- Z_{eff} -schattingen, met de bedoeling vast te stellen of de continuüm- Z_{eff} in het algemeen wordt overschat, en om een gemiddelde schaalfactor te bepalen voor de continuüm- en CX-data, en dit in verschillende ontladingsscenario's.

Ten slotte wordt een model voorgesteld voor de directe schatting van Z_{eff} -profielen, consistent met de lijngeïntegreerde remstralingsemissiviteit, de lijn-geïntegreerde elektronendichtheid en de lokale onzuiverheidsdichtheid uit CX. Het profiel wordt uitgedrukt in termen van een spline-basis. Om systematische effecten te modelleren wordt een gemeenschappelijke schaalfactor geïntroduceerd voor alle kanalen van de remstralingsdiagnostiek, en een andere factor voor alle zichtlijnen van CXS. Dit model functioneert goed op artificiële data, maar moet nog uitgetest worden op reële metingen.

Het zou waardevol zijn de Bayesiaanse modellen die geïntroduceerd worden in dit werk te verfijnen, door de gedetailleerde modellering van alle belangrijkste onzekerheden. Dit zou ook heel nuttig zijn voor het schatten van de onzuiverheidsdichtheden uit CX op zich, meer bepaald voor het bepalen van heliumconcentraties. Dit is uiterst belangrijk voor de studie van de uitstoot van helium in fusiereactoren. Daarnaast kunnen een gevoeligheidsanalyse en Bayesiaanse Diagnostische Design leiden tot een geoptimaliseerd ontwerp van de CXS- (en remstralings-) diagnostiek. Op dit moment is dat met name erg belangrijk aan TEXTOR, waar recent een nieuw CXS-systeem werd geïnstalleerd, dat een pilootexperiment is voor de actieve-bundel-diagnostiek op ITER.

English summary

1 The world energy problem

In order to sustain the standard of living of the average inhabitant of a present-day industrialized country, a significant amount of energy is required. Currently most of the energy is generated from the burning of fossil fuels and from nuclear fission. However, the majority of the reserves of both fossil fuels and uranium will be depleted within 50 years (oil, uranium) to 250 years (coal). Moreover, the burning of fossil fuels is causing serious environmental problems, e.g. due to the exhaust of greenhouse gases. There appears to be a general consensus that this is causing global climatological changes, which may even be irreversible. On the other hand, the safe handling and storage of nuclear waste from fission is a complex matter.

Several renewable energy sources exist, such as solar and wind power, but on average these methods require an enormous land use, and it is now already clear that they can only complement existing and future clean energy sources.

2 Controlled thermonuclear fusion

A possible solution is offered by *controlled thermonuclear fusion*. Nuclear fusion is the process that powers the stars, and it refers to the reaction of fusing atomic nuclei, thereby releasing the equivalent energy of the mass deficit. The research on controlled fusion on Earth aims at power production based on nuclear fusion. To this end, the fusion reaction between a deuterium (D) and a tritium (T) nucleus is considered, resulting in a helium nucleus and a neutron; the latter carrying most of the fusion energy. Due to the Coulomb barrier between two nuclei, this requires very high energies. Taking into account quantum mechanical tunneling, this is of the order of 10 keV, the equivalent of 100 million degrees centigrade. At such energies, all D and T nuclei

are ionized. The ensemble of fuel ions and electrons is called a *plasma*, which is the fourth state of matter.

The confinement of (experimental) fusion plasmas can not be done using a material wall. Instead, the plasma is trapped inside a magnetic field, the most developed configuration being the torus shaped *tokamak*. The magnetic field in a tokamak is a superposition of a toroidal field, created by toroidal field coils, and a poloidal field generated by a large inductive plasma current (hundreds of kA to MA). The charged plasma particles gyrate around the field lines, and are thus, to a certain extent, confined to the tokamak vacuum vessel. Nevertheless, there is also a radial transport of particles and energy through diffusion, convection and radiation processes, resulting in a finite particle and energy confinement time. In order to achieve thermonuclear conditions, the fuel density, temperature and confinement time should be sufficiently elevated at the same time. The plasma is heated through several techniques: ohmic heating by the plasma current, as well as auxiliary heating by the injection of energetic hydrogenic atoms (Neutral Beam Injection) and by the resonant absorption of radio frequency waves. Eventually, the power loss in the plasma must be compensated by alpha particle heating, at which point no more auxiliary plasma heating is required. This event is called *ignition*.

The present work was conducted partly at the TEXTOR tokamak (Institut für Plasmaphysik, Forschungszentrum Jülich, Germany) and JET (EFDA, Culham, UK), the world's largest tokamak. A substantial part of the worldwide efforts on tokamak research is concentrated on the next-step device ITER. The aim of ITER is to demonstrate the scientific and technical feasibility of nuclear fusion power for peaceful purposes. The actual construction phase of ITER is planned to commence in the course of 2008, while the first ITER plasma is expected in 2016. ITER will be the single step between present-day tokamaks and the first fusion reactor DEMO.

3 Plasma impurities

Although in the course of the past decades considerable progress has been made in the research for controlled nuclear fusion, several issues still inhibit the realization of thermonuclear conditions. One of the primary difficulties lies in the control of non-fuel ions, called *impurities* in the plasma. Impurities are released from plasma-facing components through various plasma-wall interaction processes, such as sputtering.

Ionised impurities may propagate towards the core plasma, where they lead to dilution of the fuel, while degrading the energy confinement through bremsstrahlung. These are the primary harmful properties of impurities, which, at reactor parameters, can prevent ignition. On the other hand, near the plasma periphery atomic line radiation by impurities can decrease the plasma-wall interaction, without deteriorating the energy confinement in the core plasma. Light impurities can even be injected with this purpose. Together with the possibilities of impurities for plasma diagnosis, this constitutes one of the main beneficial effects of impurities.

To characterize the behaviour and deduce the consequences of impurities, it is crucial to obtain information on the impurity content in the plasma. This can be quantified by the so-called *ion effective charge* Z_{eff} , which is the key quantity in the present work. It is defined as

$$Z_{\text{eff}} = \frac{\sum_i n_i Z_i^2}{\sum_i n_i Z_i}, \quad (2)$$

where n_i is the density for impurity (or hydrogenic) species i , and Z_i its charge. Z_{eff} is a local measure for the impurity concentration, averaged over all impurities and can be related directly to the electrical plasma resistivity. For ITER a Z_{eff} value of 1.8 is foreseen. Depending on the discharge scenario, this number may only vary by ± 0.2 .

4 Experimental determination of Z_{eff}

Several methods exist for the derivation of Z_{eff} . The present work deals with the calculation of Z_{eff} from bremsstrahlung spectroscopy on the one hand, and from Charge Exchange Spectroscopy on the other hand. The bremsstrahlung in a plasma predominantly originates from free-free transitions of an electron in the electric field of an ion. Once the electron density n_e and electron temperature T_e are known, the local bremsstrahlung emissivity ϵ_{ff} (usually in $\text{W}/\text{cm}^3/\text{sr}/\text{\AA}$) is directly proportional to Z_{eff} :

$$\epsilon_{\text{ff}} \sim \bar{g}_{\text{ff}}(T_e) \frac{n_e^2 Z_{\text{eff}}}{\sqrt{T_e}},$$

where \bar{g}_{ff} is the Maxwell-averaged Gaunt factor, incorporating all quantum mechanical effects. A bremsstrahlung diagnostic observes bremsstrahlung emissivity line-integrals, and a radial profile for the local emissivity has to be derived through an inversion procedure. Provided also n_e and T_e profiles are available, a Z_{eff} profile can be calculated.

At the typical energies prevailing in core tokamak plasmas, most ions are fully ionized, rendering direct impurity spectroscopy impossible. However, the energetic hydrogenic atoms from a neutral beam can cause a charge exchange reaction with a plasma ion. The plasma ion acquires an electron in an excited state, while the ionized beam particle leaves the beam. By observing the radiative decay of the plasma impurity ions, it is then possible to derive the impurity density, ion temperature and flow. This is the aim of *Charge Exchange Spectroscopy* (CXS). If the main plasma impurities are monitored, one can derive Z_{eff} according to (2). On both TEXTOR and JET carbon is the dominant impurity, so that often only the fully stripped carbon density is considered in the summation (2). This is the approach taken also in the present work. Each sight line of a CXS diagnostic observes a relatively well localized plasma volume defined by the beam geometry, so that no inversion procedure is needed to derive Z_{eff} profiles.

5 Inconsistency of the continuum and CX Z_{eff}

Many sources of uncertainty enter the derivation of Z_{eff} , both from bremsstrahlung spectroscopy (continuum Z_{eff}) and CXS (CX Z_{eff}). These uncertainties can be of a statistical or systematic nature, and they may lead to substantial errors on the calculated Z_{eff} values. In general, the continuum Z_{eff} is, indeed, inconsistent with the CX Z_{eff} . This is a long-standing problem in fusion diagnosis, and the reconstruction of a Z_{eff} profile that is reliable over the entire plasma cross-section is at present a real challenge. One of the purposes of the present work is the estimation of a Z_{eff} value that is consistent with both the bremsstrahlung emissivity and CX impurity density measurements.

6 The visible bremsstrahlung diagnostic on TEXTOR

A second component of this work concerns the design and commissioning of a new diagnostic for visible bremsstrahlung spectroscopy on TEXTOR. The main system consists of a set of 24 lines of sight, viewing part of a poloidal cross-section of the plasma. A viewing dump reduces the pick-up of reflections inside the vessel. The light is transported by fibre optics to a diagnostic room beyond the biological shield. A Fabry-Pérot interference filter selects a wavelength band that is free of

line radiation on TEXTOR. The light is focussed on a cooled CCD array. The CCD is usually kept at 0°C, where the typical signal-to-noise ratio for 50 ms exposure time and 8×8 binning at a read-out speed of 4.0 Mpixel/s, is 45 dB. During a typical data acquisition sequence, a series of images is taken with a sampling frequency of 20 Hz. Thus, the diagnostic succeeds in attaining an enhanced time resolution, compared with the previous system installed on TEXTOR, while at the same time it features a large number of chords. The timing of the diagnostic is controlled by a programmable timing circuit. The images are sent to a workstation for subsequent processing, yielding a set of line-integrated bremsstrahlung emissivity time traces. A dedicated software interface and a diagnostic control and data processing program were written in the frame of the present work. Smearing of CCD images is eliminated by estimating the contribution of smearing in a dark region of an image, without the need for calculating the Point Spread Function of the optical set-up. A set of fibres equipped with photomultiplier tubes has been installed as a back-up to the camera system.

The data are transferred to the TEXTOR central storage facility. A line-averaged Z_{eff} time trace is routinely calculated for a centrally viewing line of sight, and is available on-line via the TEXTOR Physics Database. Z_{eff} profiles can be reconstructed using an Abel inversion. This often results in a divergent Z_{eff} profile near the plasma boundary, which can be ascribed to non-bremsstrahlung contributions to the continuum radiation. Some tests have been performed using Tikhonov and Maximum Entropy regularized inversion, leading to a substantial reduction of the edge Z_{eff} overestimation.

The calibration of the system was performed using an integrating sphere, providing a uniform and constant light source. However, this procedure so far did not lead to a satisfactory relative (channel-to-channel) calibration. Therefore, an alternative technique was devised to assess the relative calibration, based on the requirement of consistency of line-integrated bremsstrahlung profiles under a horizontal shift of the plasma column. This calibration leads to physically acceptable Z_{eff} profiles. In addition, the method is generally applicable, providing a simple and self-consistent method for the relative calibration of also other multichannel spectroscopic diagnostics. No dedicated light source or any other specialized calibration set-up is needed, nor is the access to the vacuum vessel required. The latter is all the more important because frequent in situ calibrations will become increasingly difficult in future fusion devices.

7 Integrated Data Analysis for fusion diagnostics

The goal of plasma diagnosis is the derivation of reliable estimates for the physical quantities of interest, starting from the raw measured data. A nuclear fusion experiment generates large amounts of raw data that need to be processed in an efficient way. This is the phase of *data validation and analysis*. The process possibly involves the integration of data from heterogeneous sources. The derived physical quantities are affected by numerous sources of uncertainty, and a probabilistic analysis is appropriate. This leads into the realm of *Integrated Data Analysis* (IDA), where measurements from multiple diagnostics can be combined, possibly incorporating information on the underlying physical model as well. Compared to the efforts put in to enhance the sophistication of plasma diagnostic methods and diagnostic hardware, IDA is often a sideline activity. Nevertheless, this consistent processing of all available data can dramatically increase the reliability and robustness of physics results, and can contribute significantly to the optimization of diagnostic design. In addition, the available space for diagnostic set-ups at ITER will be restricted, so that any type of available information will have to be exploited for the assessment of physical quantities.

8 Bayesian inference

In this work, IDA is conducted in a Bayesian probabilistic framework (hence *Bayesian Integrated Data Analysis*, or BIDA). The Bayesian paradigm considers a probability as a degree of belief, not as a frequency. Bayesian probability theory is particularly well suited for the inference of the underlying parameters of a physical model, given a set of data, possibly supplemented with ‘expert’ prior knowledge on the parameters of interest. The basis of Bayesian inference is Bayes’ theorem, providing a means for inverting conditional probabilities:

$$\underbrace{p(\boldsymbol{\theta}|\boldsymbol{x}, I)}_{\text{posterior}} = \frac{\overbrace{p(\boldsymbol{x}|\boldsymbol{\theta}, I)}^{\text{likelihood}} \overbrace{p(\boldsymbol{\theta}|I)}^{\text{prior}}}{\underbrace{p(\boldsymbol{x}|I)}_{\text{evidence}}},$$

where \boldsymbol{x} represents a data vector, $\boldsymbol{\theta}$ are the parameters of interest, and I signifies any additional available information.

In a multiparameter model, in order to make inferences on individual parameters, one needs to marginalize the joint posterior distribution. This is required for the calculation of posterior moments as well. The multidimensional integrals can be very difficult to solve, even numerically, so that one has to resort to stochastic integration methods. In the current work, a *Markov Chain Monte Carlo* scheme is employed, in particular the *Metropolis-Hastings* sampling algorithm. It involves a Markov chain that converges to the target distribution. The estimates for the quantities of interest are inferred by calculating the Monte Carlo mean (*ergodic average*) of the sampled parameters.

The BIDA recipe can now be outlined as follows. First, the main sources of uncertainty in the data descriptive model have to be identified. Statistical uncertainties are modelled in the variance of an appropriate likelihood distribution, while systematic uncertainties are represented by so-called *nuisance parameters*. These are parameters that are not of interest, but which necessarily enter the forward model. Next, prior information is quantified, and a joint posterior distribution is formulated according to Bayes' theorem. Nuisance parameters are integrated out, and marginal moments for the parameters of interest can be calculated using, for example, MCMC. Finally, a *sensitivity analysis* permits the study of the impact of the various sources of uncertainty on the error bars of the quantities of interest. Together with techniques of *Bayesian Diagnostic Design*, this allows the optimization of the design of diagnostics.

9 IDA for the determination of Z_{eff}

In the present work, IDA is used for the estimation of a Z_{eff} value consistent both with measurements of bremsstrahlung emissivity and CX impurity density. This is carried out on data from JET discharges, since the JET CXS database is much more extensive than the one at TEXTOR. In addition, the TEXTOR CXS diagnostic had not been recently calibrated at the time of the work. The bremsstrahlung emissivity is determined from the baseline level of the CX spectrum. In a first stage, a Bayesian probabilistic model is formulated for the estimation of (time traces of) Z_{eff} on the magnetic axis. The input data are the local bremsstrahlung emissivity (and electron temperature) on the one hand, and the local carbon density obtained from CXS on the other hand. The model permits the estimation of the electron density n_e as well, which can be compared to the density from LIDAR Thomson scat-

tering, thus providing a means to check the validity of the calculations.

At first, only the statistical uncertainties on the data are modelled in terms of Gaussian likelihoods, while uninformative uniform priors are assumed. The analysis is performed on data from two JET discharges, and once without, once with the explicit inclusion of LIDAR measurements. The model is sufficiently simple to allow the calculation of the joint posterior on a grid of parameter values. A consistent Z_{eff} is obtained, but the validity of the result is questioned in cases where there is a large systematic discrepancy between the continuum and the CX Z_{eff} . This is confirmed by a consistency analysis.

Therefore, a second model is proposed where also systematic uncertainties are modelled, in terms of a scale factor for each of the measurements. The prior for the scale factors is a uniform distribution. Estimates for the parameters of interest are obtained via MCMC, namely a pure Metropolis sampling scheme with a Cauchy proposal distribution. The Cauchy scales are tuned for a favourable acceptance rate. The method is first tested on an artificial data set, and next on data from a number of JET discharges. A consistency analysis points out that the scaled input data are much more mutually consistent than their unscaled versions. The sampling algorithm performs well, and consistent Z_{eff} estimates can be determined, with significantly reduced error bars compared to the errors on the continuum and CX Z_{eff} . In several of the discharges under study, the estimated Z_{eff} time trace is more similar, both qualitatively and quantitatively, to the CX Z_{eff} , rather than to the continuum Z_{eff} . The scale factor for the continuum measurement is systematically far above unity, indicating an overestimation of the bremsstrahlung emissivity by an average factor of 1.4. In some discharges considered, the CX measurement, i.e. the summed impurity density, is underestimated by a factor of 0.4 to 0.8. These results suggest the establishment of a database of MCMC Z_{eff} estimates, with the aim to assess whether in general the continuum Z_{eff} is overestimated, and to determine average scaling factors for the continuum and CX data, in different discharge scenarios.

Finally, a model is presented for the direct estimation of Z_{eff} profiles consistent with line-integrated bremsstrahlung emissivity, line-integrated electron density and local CX impurity density. The profile is expressed in terms of a spline basis. To model systematic effects, a common scale factor is introduced for all channels of the bremsstrahlung diagnostic, and another factor for all sight lines of CXS. This model performs well on artificial data, but is yet to be tested on real

measurements.

It would be valuable to increase the sophistication of the Bayesian models introduced in this work, taking into account a detailed modelling of all primary uncertainties. This would also be most useful for the estimation of CX impurity densities in itself, particularly for the assessment of helium concentrations. This is of key interest to helium exhaust studies for fusion reactors. On the other hand, a sensitivity analysis and Bayesian Diagnostic Design can result in an optimized design of the CXS (and bremsstrahlung) diagnostic. This is at present particularly important on TEXTOR, where a new CXS system has been installed recently, which is a pilot experiment for the ITER active beam diagnostic.

What I am going to tell you about is what we teach our physics students in the third or fourth year of graduate school. It is my task to convince you not to turn away because you don't understand it. You see, my physics students don't understand it. That is because I don't understand it. Nobody does.

We have a habit in writing articles published in scientific journals to make the work as finished as possible, to cover up all the tracks, to not worry about the blind alleys or describe how you had the wrong idea first, and so on. So there isn't any place to publish, in a dignified manner, what you actually did in order to get to do the work.

Richard Feynman

1

Introduction

1.1 Topics of interest in this thesis and situation

The rapid use of world energy resources and the impact of energy supply on the global environment, is causing increasing concern among scientists and policymakers. These issues are all the more pressing since expectations are that the worldwide need for energy will do nothing but increase in the decades ahead. There is therefore an urgent necessity to explore alternative possibilities for energy supply that are at the same time able to meet the demands, while reducing the per-

nicious effects on the environment. A strong candidate is controlled thermonuclear fusion, which offers a prospect of a clean, safe and virtually inexhaustible power source. The realization of nuclear fusion for power production on Earth is a scientifically and technically extremely challenging enterprise. The research on controlled fusion has come a long way, and many successes have been achieved, but many difficulties still have to be tackled.

The research described in this text fits in with the search for controlled nuclear fusion via magnetic confinement of a hydrogenic plasma in a tokamak. In particular, the main interest of this work lies in the determination of a quantity that is of vital importance for the understanding of many physical processes in the plasma, namely the ion effective charge Z_{eff} . In addition, in a future fusion reactor, Z_{eff} is a number that will need to be controlled accurately, since it determines for a considerable part the performance characteristics of the reactor. Indeed, Z_{eff} is a measure for the general impurity concentration in the plasma, while impurities play a very important role in tokamak physics, through their various harmful and beneficial effects on the plasma.

Thus, it is crucial to dispose of reliable and robust measurements of Z_{eff} . There are several diagnostic methods to estimate Z_{eff} , the most popular occurring through bremsstrahlung spectroscopy on the one hand, and Charge Exchange Spectroscopy (CX) on the other hand. However, both methods are subject to various sources of uncertainty, rendering the Z_{eff} estimates from bremsstrahlung emissivity measurements in general in poor agreement with the estimates from the individual impurity concentrations determined by CX.

The main purpose of the present work is to estimate a value for the local Z_{eff} in the plasma that is consistent both with the measurements of bremsstrahlung emissivity and with the CX impurity concentration data. To this end, we reformulate the determination of Z_{eff} from the respective data sets into a joint probabilistic model, encoding the main sources of uncertainty. We employ Bayesian probability theory, which provides a clear recipe for the estimation of the underlying parameters (in this case Z_{eff}) determining a physical model, given a set of measurements, while taking into account ones prior beliefs about the nature of the parameters as well.

This strategy fits in with the concept of a Bayesian Integrated Data Analysis of fusion diagnostic data. The motivation for this approach is the recognition that the diagnosis of hot plasmas is inherently diffi-

cult, and that measured quantities are typically affected by large error bars. In order to cope with this problem, the general tendency is to apply increasingly sophisticated diagnostic methods, and to enhance the accuracy and capabilities of diagnostic hardware. While this is an absolutely necessary development in order to improve the general dependability of plasma diagnosis, it should not be neglected that there exists a whole spectrum of data analysis techniques that can also dramatically improve the quality of measured data and the reliability of derived physical quantities. One of the possibilities is given by Integrated Data Analysis, the underlying idea being the combination of heterogeneous sources of information for the estimation of a single quantity. This is of particular benefit when one is confronted with data sets that lead to contradictory results for the quantity of interest, as it is the case for the determination of Z_{eff} .

In the current work, a first approach is presented towards a Bayesian Integrated Data Analysis for the estimation of Z_{eff} from measurements of bremsstrahlung emissivity and CX impurity concentrations. The proposed model is rather uncomplicated, leaving still room for the various uncertainties to be modelled in more detail. Nevertheless, we will demonstrate that it is possible, in the context of the current model, to derive Z_{eff} values that are consistent with both the bremsstrahlung and CX data. In addition, the analysis will allow us to draw several conclusions on the accuracy of the two diagnostic methods involved.

A second component of the work described here, concerns the design and commissioning of a new, improved diagnostic for bremsstrahlung emissivity measurements in the visible range on the TEXTOR tokamak. This diagnostic is now in continuous operation, providing routine measurements of a line-averaged Z_{eff} in the TEXTOR plasmas, with the additional possibility to reconstruct radial Z_{eff} profiles for a local estimate of the impurity content.

This work has started out on the TEXTOR tokamak, and has lead to the successful installation of the visible bremsstrahlung diagnostic on that machine. However, several issues, primarily the calibrations, related to the bremsstrahlung and CX diagnostics on TEXTOR have prevented the use of TEXTOR data for the Integrated Data Analysis for Z_{eff} (see Section 3.2.5). Therefore, this analysis has been performed using data from the CX diagnostic on the JET tokamak.

1.2 Outline

Chapter 1 is this chapter, the introduction, providing also a list of publications that have arisen from the current work.

In Chapter 2, we start our discussion with a brief overview of the energy problem, some aspects of tokamak physics, and in particular a short account on impurities and Z_{eff} in magnetically confined plasmas.

Chapter 3 presents an outline of the two diagnostic methods that are used in this work to assess Z_{eff} , namely bremsstrahlung spectroscopy and Charge Exchange Spectroscopy. This includes an overview of the most important sources of uncertainty that enter the derivation of Z_{eff} from the respective measurements.

In Chapter 4, the main aspects of the new diagnostic for visible bremsstrahlung measurement and Z_{eff} determination on TEXTOR are introduced. Several issues are discussed in some detail, in particular the development of a novel technique for the determination of the relative calibration of the diagnostic, using the plasma itself as a reference source.

Next, a chapter has been included (Chapter 5) on Bayesian probability theory. This chapter also gives a short introduction on a technique that we have used to sample from (high-dimensional) intractable probability distributions, namely Markov Chain Monte Carlo.

The results from our Bayesian Integrated Data Analysis for the estimation of Z_{eff} are presented in Chapter 6. A probabilistic model is proposed for the estimation of an on-axis Z_{eff} value from the bremsstrahlung and CX measurements, modelling both statistical and systematic uncertainties. Then, another model is tested that allows the direct estimation of full Z_{eff} profiles from line-integrated bremsstrahlung and electron density data, and local CX impurity density measurements.

Finally, the general conclusions are drawn in Chapter 7 and an outlook is given towards a continuation of the current work.

1.3 Publications

The work described in this thesis has led to the publications listed below. Most of the work on the Integrated Data Analysis for the estimation of Z_{eff} , described in Chapter 6, has not been published yet, and this will be the subject of two journal publications that are in preparation.

International refereed journals (A1)

1. G. Verdoolaege, G. Telesca, E. Delabie, G. Van Oost and the TEXTOR team, ‘Design and commissioning of a new diagnostic for routine determination of Z_{eff} from visible bremsstrahlung measurements on TEXTOR’, *Rev. Sci. Instrum.* **77**, 10F310, 2006.
2. G. Telesca, K. Crombé, M. Tokar, B. Unterberg, G. Verdoolaege *et al.*, ‘Preliminary study of the influence of DED on carbon radiation and transport in the TEXTOR tokamak’, *J. Nucl. Mater.* **337–339**, 361, 2005.
3. A. Donné, M. De Bock, I. Classen, M. von Hellermann, G. Verdoolaege *et al.*, ‘Overview of core diagnostics for TEXTOR’, *Fusion Science and Technol.* **47**, 220, 2005.
4. D. Kalupin, P. Dumortier, A. Messiaen, M. Tokar, G. Verdoolaege *et al.*, ‘Impact of hydrogen fuelling on confinement properties in radiative improved mode’, *Plasma Phys. Control. Fusion* **45**, 1501, 2003.

National journals (A4)

1. G. Van Oost, G. Verdoolaege and J. Ongena, ‘Kernfusie: energiebron voor de toekomst’, *Tijdschrift VeLeWe (Vereniging Ler-aars Wetenschappen)*, 2002.

Proceedings of international conferences (C1)

1. G. Verdoolaege, M. von Hellermann, R. Jaspers, M. Ichir and G. Van Oost, ‘Integrated Bayesian estimation of Z_{eff} in the TEXTOR tokamak from bremsstrahlung and CX impurity density measurements’, *Proceedings of the 26th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Paris, 2006.
2. G. Verdoolaege, G. Telesca, G. Van Oost and the TEXTOR team, ‘Optimization of the relative calibration for a visible bremsstrahlung Z_{eff} diagnostic on TEXTOR via requirements of profile consistency’, *Proceedings of the 33rd EPS Conference on Plasma Physics and Controlled Fusion*, Rome, P-2.165, 2006.

3. G. Telesca, G. Verdoolaege, W. Biel, R. Jaspers, M. Lehen, Y. Liang, B. Unterberg, G. Van Oost, R. Zagorski, 'Investigation of carbon screening on TEXTOR with Dynamic Ergodic Divertor in 6/2 mode', Proceedings of the 33rd EPS Conference on Plasma Physics and Controlled Fusion, Rome, P-2.157, 2006.
4. G. Verdoolaege, G. Telesca, G. Van Oost and the TEXTOR team, 'Integrated Bayesian estimation of Z_{eff} on TEXTOR from Bremsstrahlung and CX Impurity Density Measurements', Proceedings of the 4th Workshop on Fusion Data Processing, Validation and Analysis, Jülich, 2006.
5. G. Verdoolaege, G. Van Oost, M. Ichir, A. Mohammad-Djafari and M. von Hellermann, 'Bayesian estimation of Z_{eff} in the TEXTOR tokamak from visible bremsstrahlung and CXRS measurements', in K.H. Knuth, A.E. Abbas, R.D. Morris and J.P. Castle, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Vol. Conf. Proc. 803, p. 456, AIP, Melville, NY, 2005.
6. G. Telesca, G. Verdoolaege, K. Crombé, B. Unterberg and G. Van Oost, 'Screening and radiation efficiency of carbon with Dynamic Ergodic Divertor on TEXTOR', Proceedings of the 32nd EPS Conference on Plasma Physics and Controlled Fusion, Tarragona, P-1.020, 2005.
7. B. Unterberg, S. Abdullaev, W. Biel, M. De Bock, G. Verdoolaege *et al.*, 'Overview of transport studies with the Dynamic Ergodic Divertor in the tokamak TEXTOR', Proceedings of the 2nd Workshop on Stochasticity in Fusion Plasmas, Jülich, 2005.
8. G. Verdoolaege, G. Telesca and G. Van Oost, 'Reconstruction of Z_{eff} profiles at TEXTOR through Bayesian source separation', Proceedings of the 12th International Congress on Plasma Physics, Nice, CCSD-00001802, 2004.
9. G. Verdoolaege, G. Telesca and G. Van Oost, 'Reconstruction of Z_{eff} Profiles at the TEXTOR Tokamak Through Bayesian Source Separation', in R. Fischer, R. Preuss and U. von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Vol. Conf. Proc. 735, p. 344, AIP, Melville, NY, 2004.

10. D. Kalupin, P. Dumortier, A. Messiaen, M. Tokar, G. Verdoolaege *et al.*, ‘Dependence of confinement on fuel isotope in impurity seeded plasmas’, Proceedings of the 30th EPS Conference on Plasma Physics and Controlled Fusion, St. Petersburg, P-2.136, 2003.
11. G. Verdoolaege, G. Telesca and G. Van Oost, ‘Neural network methods for radial profile reconstruction for Z_{eff} from bremsstrahlung data on the TEXTOR tokamak’, in D. Ruan, P. D’hondt and E. Kerre, editors, Computational Intelligent Systems for Applied Research, Proceedings of the 5th International FLINS Conference, p. 615, World Scientific, Singapore, 2002.
12. G. Verdoolaege, G. Telesca, G. Van Oost and G. Van Den Berge, Reconstruction of the edge Z_{eff} profile from bremsstrahlung data via extensions of Independent Component Analysis on TEXTOR, Proceedings of the 29th EPS Conference on Plasma Physics and Controlled Fusion, Montreux, P-2.060, 2002.

The essential fact which emerges [...] is that the three smallest and most active reservoirs [of carbon in the global carbon cycle], the atmosphere, the plants and the soil, are all of roughly the same size. This means that large human disturbance of any one of these reservoirs will have large effects on all three. We cannot hope either to understand or to manage the carbon in the atmosphere unless we understand and manage the trees and the soil too.

Freeman Dyson

2

Controlled thermonuclear fusion and tokamak physics

2.1 The energy problem

The modern society as we know it depends more than ever on the availability of reliable sources of energy. In some parts of the world, like in Western Europe, the demand is higher than in other regions, like in many developing countries, and the supply of energy therefore occurs on different scales. Areas with a high population density or a strongly developed economy, require a large energy provision, often centralized

Country	Per capita consumption (W)
Canada	13.200
Norway	13.000
USA	11.200
Japan	5700
Europe (West and East)	4800
Former Soviet Union	4000
China	990
India	370
Developing countries	100 – 1000
World	2100

Table 2.1: Average annual per capita total primary power consumption for selected countries (figures date from 1995) [4].

for the supply of cities and industry. Table 2.1 shows the average annual per capita primary power consumption for a selection of countries, showing clearly the dominant position of the Western countries, Japan and also the countries part of the former Soviet Union. Currently, the global annual energy consumption amounts to more than 12 TWyr. However, taking into account the rising energy demands as well as the predicted rise of world population to 10 billion in the next 50 years [6], a quite realistic estimate predicts world energy consumption of about 30 TWyr around 2050 [4]. A question that rises in this context, is whether it is possible—and for how long—to satisfy these enormous energy demands. Table 2.2 shows the reserves for the most used fuels that have been proved to exist, and for which the exploitation is economically feasible [4]. The mentioned time spans are not long—indeed the current young generation will still live to see the actual depletion of oil, natural gas and uranium. Although nuclear power generation might be sustained for a much longer time via the *breeding* of fuel, the associated safety and environmental issues might be difficult to cope with [7]. However, new reactor concepts could change this situation [8]. Although there is substantial uncertainty in the numbers in Table 2.2, the data do suggest that gradually a change in energy policy is needed, and that alternatives for the main current fuels should be looked for and exploited. The rapid use of world energy sources is causing also other problems. Firstly, fuel scarcity is already now leading to political

Fuel	Proved recoverable reserves	Years of use at the current rate of con- sumption
Coal	1.0×10^{12} tons	270
Crude oil	950×10^9 barrels	40 – 50
Natural gas	120×10^{12} m ³	60 – 70
Uranium	2.0×10^6 tons	40 – 50 (2400 – 3000 if breeder technology is employed)

Table 2.2: Estimated maximum years of use of different fuels at the current rate of consumption [4].

instabilities (1970's energy crisis, Gulf War, war in Chechenia, etc.). Secondly, some of the fuels that are now used for energy supply, are at the same time valuable resources in other industrial branches, like oil in the chemical and pharmaceutical industry. Thirdly, and this is perhaps the most worrisome problem, there are many environmental issues associated with the present ways of energy generation. The massive burning of fossil fuels leads to the release of enormous quantities of carbon dioxide in the Earth's atmosphere. Figure 2.1 shows the steep increase of CO₂ concentration in the atmosphere since the beginning of the Industrial Revolution at the end of the 18th century. It is still not clear what the possible consequences are of such sudden large changes in the composition of the atmosphere, but there seems to be at least general agreement among specialist that, since CO₂ is a greenhouse gas, the average environmental temperature is rising as a result of CO₂ emission by human activity (see Ref. [4] and references therein). This temperature rise in turn may have several undesirable and possibly irreversible effects on the global environment. On the other hand, nuclear energy generation is a source of highly radioactive waste, and its safe handling and storage is a very complex matter and the subject of many environmental debates. Nevertheless, the amount of waste produced by fission is, for the same amount of generated power, much smaller than compared to burning fossil fuels. Moreover, as mentioned before, new reactor technologies might very well improve the safety and waste control aspects of fission. In this context the issues can be

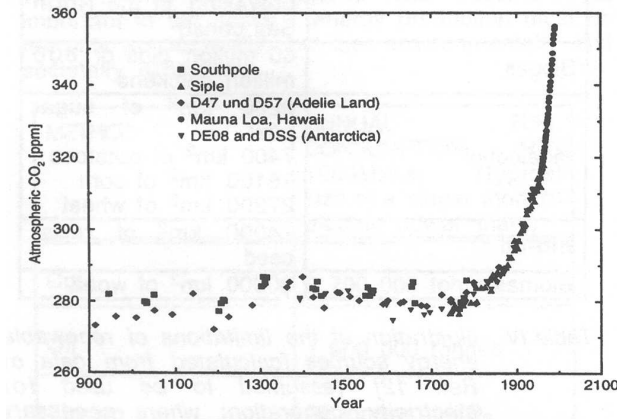


Figure 2.1: Evolution of the CO_2 concentration in the atmosphere (in ppm) during the last 1000 years [1].

mentioned associated with the proliferation of nuclear waste.

Although a cautious measure would be an immediate and drastic reduction of energy generation from fossil fuels, this may in practice be very difficult to realize. None of the possible alternatives, except for fission, is presently mature enough for large scale energy production, while also fission has, as mentioned, a number of disadvantages. The situation is even worsened by the dramatic decrease of energy research budgets in the last decades and the liberalization of the energy market in Europe, which may well lead to an increase in energy consumption. Also, there are tremendous economic and political powers trying to maintain the current situation (see e.g. Ref. [9]).

Regarding the often mentioned renewable energy sources, it should be clear that they are not real alternatives to energy generation from fission and fossil fuels. Rather, they would complement existing and future cleaner energy sources. Indeed, Table 2.3 shows for several types of renewable energy sources the respective investments needed to produce 1 GW of electricity—the typical amount of power from a single modern electric power plant. It turns out that the land use needed for the energy supply by renewables of the average industrialized country would be unrealistically high. So, although it makes certainly sense to try to exploit every possible alternative to fossil fuel combustion, one should bear in mind the limited prospects for renewables [4]. With respect to the current situation, Figure 2.2 gives an overview of the

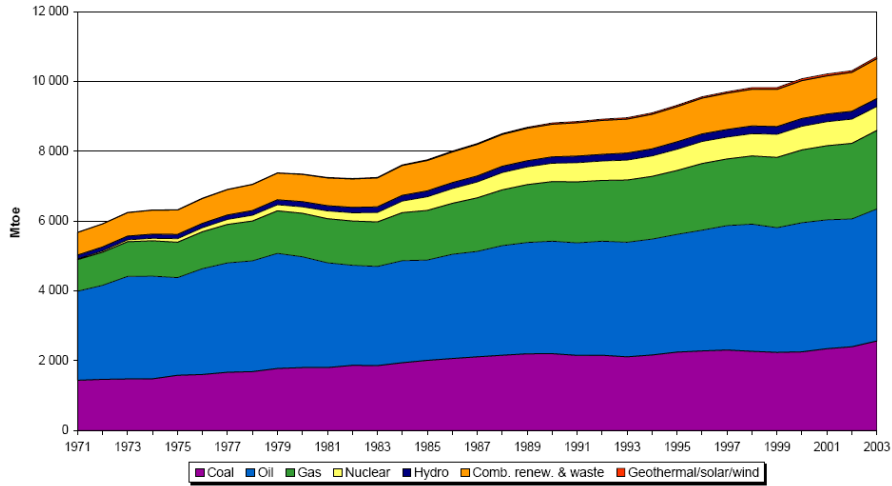


Figure 2.2: Annual world energy generation, distributed per energy type (expressed in Million Tons Oil Equivalent), showing also the rise of energy needs during the last decades.

annual world total energy generation, split up between the most used energy types [10]. The rise in energy demand can clearly be seen, as well as the small relative contribution of renewable energy sources.

Next to fission and renewables, a third candidate for future power generation is investigated. This option is nuclear fusion. It is the least developed of the three, but it holds the promise of being a safe, inexhaustible and clean method for energy generation [11], [12]. The context of the work described in this text is the quest for energy production by controlled nuclear fusion. The purpose of the remainder of this chapter is to make clear the relation of the topics studied in this thesis with the current research on controlled fusion. We will give a very brief overview of some of the aspects related to tokamak physics, focussing on the topics that are of the most relevance to the current work. For a general overview on tokamak physics, see e.g. Refs. [13], [2], [14], [15] and [16].

2.2 Controlled thermonuclear fusion

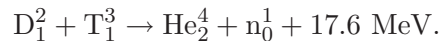
The term *thermonuclear fusion* refers to the fusion of two atomic nuclei through their thermal energy. The resulting product has a lower total

Method	Needed surface
Photovoltaic panels	about 100 km ² in Middle Europe (10% efficiency assumed)
Windmills	6660 mills of 150 kW (20 m rotor blades and for the average wind speed at the North Sea coast)
Biogas	60 million pigs or 800 million chickens
Bioalcohol	6200 km ² of sugar beet, or 7400 km ² of potatoes, or 16.100 km ² of corn, or 27.200 km ² of wheat
Bio-oil	24.000 km ² of rapeseed
Biomass	30.000 km ² of wood

Table 2.3: Investments needed to produce 1 GW of electricity for several types of renewables [4]. Surfaces can be compared to the total surface of Belgium: 32.545 km².

mass than the sum of the two original nuclei, and the mass deficit Δm is converted into energy through $E = \Delta mc^2$. This is the energy source that powers the stars, and the ultimate purpose of the research on controlled thermonuclear fusion is to produce energy by fusion in a power plant on Earth.

Although many nuclear fusion reactions exist, only the fusion of a deuterium (D) with a tritium (T) nucleus appears to be scientifically, technically and economically feasible for power production at this point. The reason is that the nuclei are repelled from each other by the Coulomb force resulting from their electrostatic charge. Hence, the energy of a collision has to be high enough to overcome this repulsion, and thus allow the nuclei to fuse through the Strong Nuclear Force. The cross-section of the D-T reaction is sufficiently large at thermal energies that are within reach in a laboratory (Figure 2.3). In a D-T fusion reaction, a deuterium and tritium nucleus fuse, to form an alpha particle and a neutron, according to



The released energy comes in the form of kinetic energy for the alpha particle (3.5 MeV) and, by far the largest part, for the neutron: 14.1 MeV. The release of energy is enormous, considering that the mass

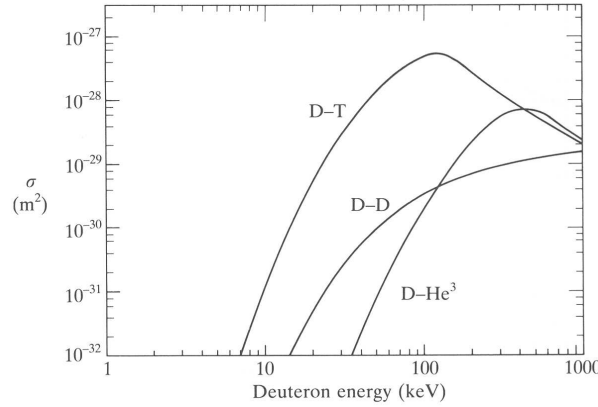


Figure 2.3: Cross-sections of the reactions D-T, D-He³ and the total cross-section of the two possible D-D reactions.

deficit for this reactions is only 0.01875 times the mass of the proton. In more practical terms, if one would start from 1 kg of D-T fuel, the fusion of all its nuclei according to the D-T reaction would release 10^8 kWh of energy, which is the amount of energy produced by a 1 GW (electrical) power station for a day.

Now, in order to reach thermonuclear energies, one first has to heat the D-T fuel to the order of 10 keV, which is, in thermal equilibrium, the equivalent of 100 million degrees centigrade—hotter than the centre of the Sun. At such temperatures, all fuel ions are ionized, and the electrostatic charge of the nuclei is neutralized by an equal number of electrons. The resulting neutral ionized gas is called a *plasma*. Plasmas can also exist at much lower temperatures and most of the visible matter in the universe is in the plasma state.

A gas at this high a temperature can not be confined by a material wall. Another means of confinement is offered by the so-called *tokamak concept*, wherein the plasma particles are confined by a magnetic field. The plasma conditions required for a sufficient fraction of fuel to fuse come down to a simultaneous sufficiently high ion temperature, ion density and energy confinement time. The required values for these three plasma parameters have all been reached already in nuclear fusion research, but unfortunately not yet in the same plasma. However, the progress in the last decades has been enormous, as illustrated by Figure 2.4. Eventually, just as the energy contained in the logs of a

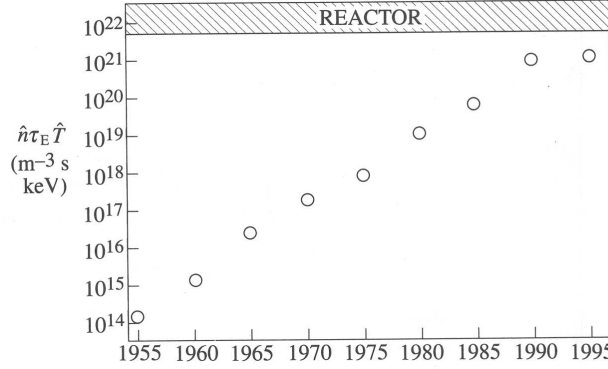


Figure 2.4: The progress in fusion research towards reactor conditions, as measured by the product of peak ion density \hat{n} , energy confinement time τ_E and peak ion temperature \hat{T} .

wood fire is high enough to set freshly added logs on fire, the energy of the alpha particles should at a certain point be sufficient to sustain the fusion reactions, despite of the inevitable energy losses of various kinds. This event is called *ignition*.

Another scheme for magnetic confinement fusion is the *stellarator*, wherein the magnetic field is entirely generated by magnetic field coils of a special shape. We will not treat stellarators here any further. Neither will we talk about yet another line of fusion research, namely *Inertial Confinement Fusion*.

2.3 Tokamaks

As mentioned before, in a tokamak the plasma is confined by a magnetic field¹. A tokamak is an example of a low pressure gas discharge tube in a toroidal configuration, in which the principal field is the toroidal field B_ϕ (see Figure 2.5), generated by a set of toroidal field coils. Since the plasma particles are charged, they gyrate around the field lines, and are thus confined in the direction perpendicular to the magnetic field. However, if there were only a toroidal field component, plasma particle drifts would prevent the confinement. Therefore, a second, poloidal

¹The word tokamak is a Russian acronym for **toroidalnaya kamera** and **magnitnaya katushka**, meaning ‘toroidal chamber’ and ‘magnetic coil’.

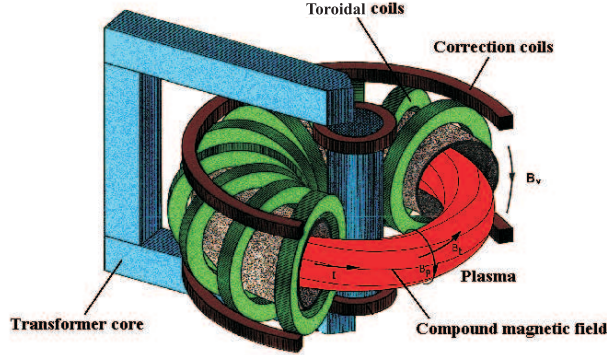


Figure 2.5: The principle of magnetic confinement in a tokamak.

component B_θ is added to the toroidal field. This component is created by generating a large electrical current (order hundreds of kA to MA) in the plasma, which is a very good conductor. The sum of the large toroidal component, and in comparison small poloidal component, is a helical magnetic field. The plasma current is generated inductively, whereby the plasma acts as the secondary of a transformer. An iron transformer core is often used and is also sketched in Figure 2.5. On the TEXTOR tokamak, the magnetic flux surfaces are nested toroids, but other shapes are possible, such as the D-like shape on the JET tokamak (see Figure 2.16)². The inboard side of the tokamak, close to the vertical axis, and where the toroidal field is the highest, is commonly called the *High Field Side* (HFS), the outboard side is the *Low Field Side* (LFS).

2.4 Plasma equilibrium

In plasmas of fusion interest, the ions and electrons have near-Maxwellian distributions. Hence the plasma pressure can be written as $p = nT$, where the temperature is, as usual in fusion science, expressed in energy units, and $n = n_e + n_i$ is the sum of the electron and ion number density. The basic condition for the plasma to be in equilibrium is that the force due to the plasma pressure p balances the force from the magnetic field

²The TEXTOR and JET tokamaks are described in more details in Sections 2.12 and 2.13, respectively.

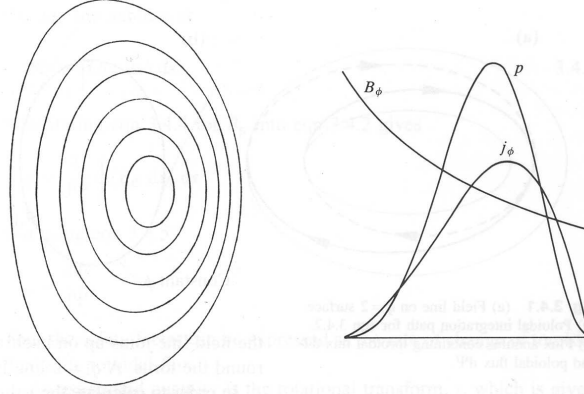


Figure 2.6: Equilibrium flux surfaces and a plot of the toroidal current density j_ϕ , the plasma pressure and the toroidal magnetic field on the midplane.

\mathbf{B} , i.e.

$$\mathbf{j} \times \mathbf{B} = \nabla p, \quad (2.1)$$

with \mathbf{j} the plasma current density. It follows that $\mathbf{B} \cdot \nabla p = 0$, so that the magnetic field lines lie entirely in surfaces of constant pressure. Since the transport across the magnetic field is much slower than the transport along the field, the plasma properties can as a good approximation be taken as constant on a flux surface.

Introducing the *poloidal magnetic flux function* ψ , which determines the poloidal flux lying within each magnetic surface, one can write (2.1) as a differential equation in ψ , the so-called *Grad-Shafranov equation*. Figure 2.6 shows the flux surfaces and profiles resulting from numerical solution of this equation for a typical case. It can be seen that the centres of the flux surfaces are displaced with respect to the centre of the tokamak vacuum vessel. This displacement is called the *Shafranov shift*.

2.5 Plasma confinement

To achieve thermonuclear conditions in a tokamak, it is necessary to confine the fuel ions for a sufficient time, in order to have a significant fraction of them to fuse. On the other hand, particles should not be

confined so long that the burned fuel, namely the alpha particles, becomes a major fraction of the plasma, leading to a strong dilution of the fuel. In addition, also energy must be confined long enough for the plasma to reach thermonuclear temperatures.

The confinement in a tokamak is not total (and can never be), although ions typically travel a distance a million times the dimensions of the vessel before hitting the wall. The efficiency of confinement by the magnetic field is commonly expressed by the *plasma- β* :

$$\beta \equiv \frac{p}{B^2/2\mu_0}.$$

In the absence of plasma instabilities, the confinement is governed by Coulomb collisions, and the transport of particles and energy can be described in terms of diffusion and convection processes (taking into account also the plasma resistivity). This allows particles to step across the magnetic field, thus deteriorating the confinement. The energy confinement is limited by thermal conduction and convection, and by radiation. The collisional transport in a torus is known as *neoclassical transport* (as opposed to classical transport in a cylinder), and this transport can be calculated. Unfortunately, in reality the transport in a plasma is much higher due to plasma instabilities, and so far no model has been able to adequately describe this so-called *anomalous transport*. Particle transport is treated in a bit more detail in Section 2.8.3, together with impurity transport.

2.6 Plasma-surface interaction

In a tokamak discharge, plasma particles inevitably escape the confinement at a certain moment, and eventually strike the material boundary surface. When a hydrogenic ion or atom, or an electron reaches a solid surface, there are three possibilities:

- (a) the particle may be directly back-scattered or reflected into the plasma, with some fraction of the impact energy,
- (b) the particle may become implanted in the surface, undergoing thermal relaxation and eventually being released again under some form,
- (c) as (b), but the particle may remain trapped in the solid for an extended period of time, or even permanently.

If the incoming particle is a hydrogenic ion, it will extract an electron from the solid as it enters, and become neutral. In addition, electrons can also stick to solid surfaces. A solid surface therefore acts as a sink for plasma, although it is not a mass sink, since most of the particles are subsequently released again in a neutral atom or molecule. The steady-state condition whereby plasma ions and electrons are captured by the surface at the same rate as recombined neutrals enter the plasma, is called *recycling*.

The impact of a particle with the wall, which is often made of Carbon Fibre Composite(s) (CFC) or graphite, can induce several processes of *plasma-surface interaction* that damage the wall [17]. For incoming ions or atoms, the principal mechanisms are physical and chemical sputtering. Physical sputtering occurs when the incoming particle is sufficiently energetic to transfer enough momentum to an atom in the solid lattice to eject it. Thus, gradually particles from the wall are injected into the plasma. Chemical sputtering by hydrogenic ions or atoms refers to the process where the chemical potential energy of the incoming ion or atom is available to break e.g. C–C bonds, and create C–H bonds. This leads to the formation of compounds such as CH₄, which are released from the surface into the plasma. The incoming particle does not even need to be very energetic to cause chemical sputtering. Next to sputtering, evaporation of plasma-facing structural components is another important source of wall particles in the plasma. Furthermore, electrons can also be released from a solid surface into the plasma by impact of electrons, ions, photons, excited neutrals, etc.

2.7 Limiter and divertor tokamaks

In order to reduce the damaging of the material walls in a tokamak, the plasma-surface interaction is concentrated on relatively small surfaces (order 1 m²) that are specially suited for high power loads. Two configurations are commonly used, the *limiter* and *divertor* tokamak.

The principle of a limiter is to introduce a material surface at a certain position inside the tokamak vessel, protruding inward from the vessel walls. A few possibilities are depicted in Figure 2.7. The concept of a limiter is based on the very fast transport of particles along the magnetic field, compared to the cross-field transport. Then, as soon as a particle reaches the radial position of the limiter surface, it will quickly hit the limiter surface and be released again as a thermalized

neutral. Thus, the plasma is limited to a certain radius, which is slightly larger than the limiter radius a , the *minor radius* of the tokamak. The minor radius for a non-circular plasma cross-section is defined as the minor radius measured along the midplane. The *major radius* R_0 is the distance between the vertical axis of the torus and the geometric central axis of the vacuum vessel. This is shown in Figure 2.8. The thin annular region outboard of the limiter radius is called the *Scrape-Off Layer*, or SOL. The magnetic field lines striking the limiter are called *open*, the others are named *closed*. The *Last Closed Flux Surface* (LCFS) is the last flux surface, when going outwards from the plasma centre, that does not touch a solid surface. The plasma inside the LCFS is called the *main* or *confined* plasma. If the limiter is pumped, involving a pump duct very near to the LCFS, in addition the neutrals can be pumped away. This can be useful for a better control over the plasma density.

In a divertor configuration, an external conductor carrying a current I_D in the same direction as the plasma current I_p , produces a change in the magnetic configuration, as shown in Figures 2.9 and 2.10. In a poloidal plane, the magnetic field lines make a figure eight shape. At some point the total poloidal field vanishes, which is a magnetic *X-point*. The magnetic flux surface passing through the X-point is called the magnetic *separatrix*, which is also the LCFS. Inside the separatrix, surrounding the plasma current, exists the main plasma. The region below the X-point and inside the separatrix is called the *private plasma*. As in the case of a limiter machine, a plasma sink is introduced by a solid material surface, cutting through the magnetic field lines surrounding the divertor conductor, see Figure 2.10. These are called the *divertor target plates*. A divertor machine may also contain additional limiters. The divertor can be equipped with pumps, and this is of special importance to the pumping of helium, originating from the fusion reactions, because helium has the ability to recycle naturally. A divertor configuration has a number of advantages over a limiter configuration, the most important being an improved energy confinement and better helium pumping.

Ideally, the plasma would only make contact with the limiters and divertor targets. In practice this is not always achieved, and particles may also be released from other parts of the vacuum vessel. Even if there were no plasma-wall contact, neutrals resulting from charge exchange (see Section 3.2.2) would still bombard all vessel surfaces,

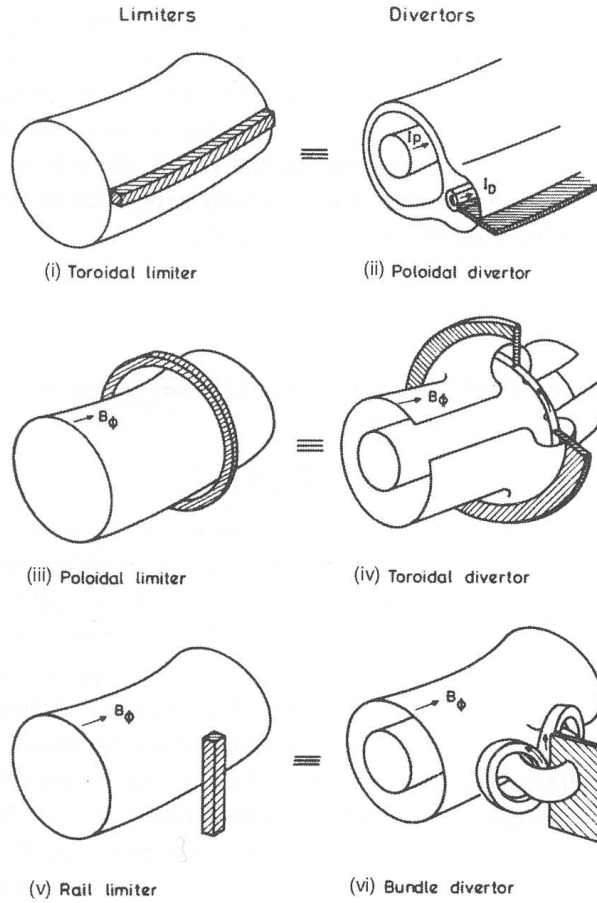


Figure 2.7: Several possible limiter and divertor configurations [2]: (i) toroidal limiter, which is analogous to the poloidal divertor in (ii), the poloidal limiter (iii) and toroidal divertor (iv), the rail limiter (v) and bundle divertor (vi).

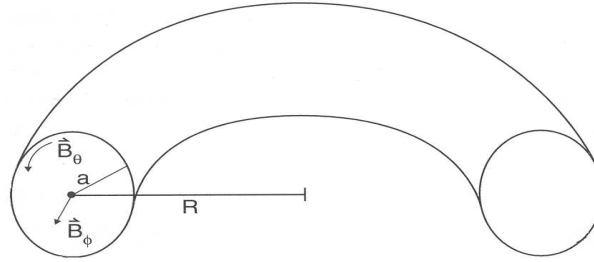


Figure 2.8: Schematic showing the minor radius a and major radius R_0 for a tokamak with a circular poloidal cross-section.

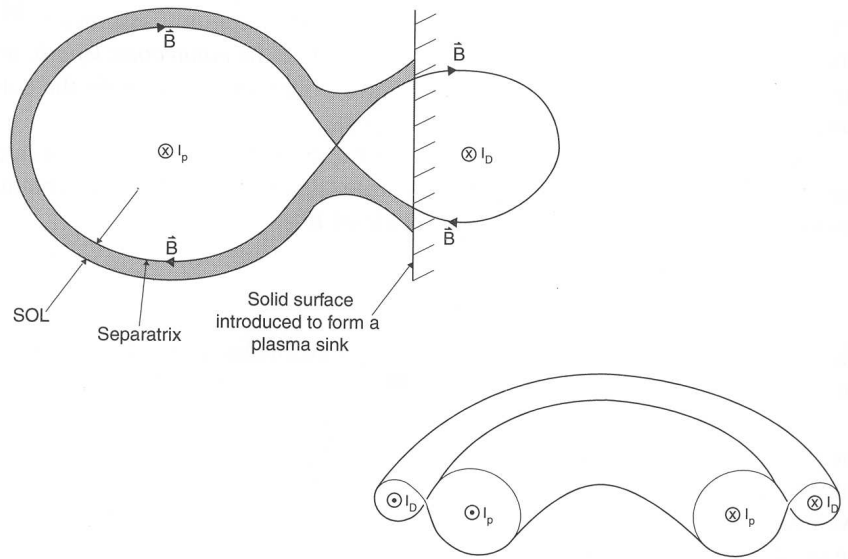


Figure 2.9: Schematics of a divertor configuration, where the poloidal field \vec{B}_θ is diverted by the divertor coil, internal or external to the vacuum vessel.

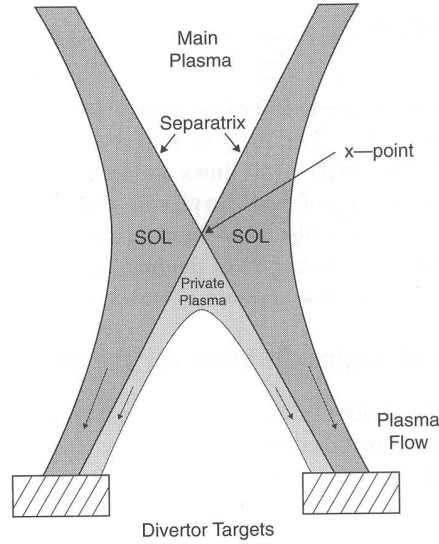


Figure 2.10: The divertor SOL, showing also the divertor target plates.

causing chemical sputtering.

2.8 Plasma impurities

Ideally, in a fusion plasma only hydrogenic fuel ions would be present. In reality, plasma impurities, i.e. non-fuel ions, are unavoidable and play even a major role. To begin with, the helium ash from the nuclear reactions is of course present. Then, as we have already seen, due to plasma-surface interaction, atoms and molecules from the solid structural components surrounding the plasma, are released into the plasma. For a limiter machine, these components are primarily the limiters, and for a divertor machine they are mainly the divertor target plates and additional limiters. Therefore, the material that makes up the limiter surfaces and divertor target plates determines which impurity becomes dominant.

Incidentally, impurities that are injected deliberately (see below) are called *extrinsic* impurities, as opposed to the ever present *intrinsic impurities*, such as carbon from the wall components in some machines. A review paper on plasma impurities can be found under Ref. [18].

2.8.1 Harmful and beneficial effects of impurities

Plasma impurities have a number of properties that have harmful consequences, but we will only consider the most important here. Firstly, impurities cause a loss of power from the plasma by bremsstrahlung (see Section 3.1.1) and atomic line radiation. This is also true for the hydrogenic fuel species, but to a far lesser degree because of their lower atomic number. Depending on their atomic number, and ionization and excitation potential, different impurity species will radiate in different plasma regions. Generally, the power is preferentially radiated from hotter plasma regions when the atomic number Z increases. In addition, the radiated power per impurity atom is also a function of Z . The bremsstrahlung emissivity is proportional to Z^2 (Section 3.1.1), while the energy loss from line radiation is proportional to $Z^3 - Z^4$. Moreover, impurities dilute the burning fuel. Both the effects of power radiation and fuel dilution, when extrapolated to reactor conditions, would prevent the ignition of the plasma. For this, 3% of low- Z elements (like carbon or oxygen), 1% of intermediate- Z elements (such as iron) or 0.1% of high- Z elements (like tungsten) would already be sufficient [19]. Some stability issues related to the impurity content are also of importance in present-day fusion machines. Finally, if a plasma impurity, after injection into the plasma, is ionized and eventually strikes again a solid surface, it may give rise to *self-sputtering*. This process is particularly damaging for the wall because momentum transfer between similar masses is very effective, as compared to sputtering by lighter particles.

On the other hand, impurities also have a number of beneficial effects. The most important of these is the power loss that occurs primarily in the SOL or near the periphery of the main plasma, without deteriorating the energy confinement in the core plasma. The advantage is that this radiated power loss diminishes the power load on the plasma-wetted areas. This natural effect can be enhanced by the deliberate injection (or *puffing*) of low- Z impurity gases such as nitrogen and neon, giving rise to a so-called *cold plasma mantle* at the periphery of the main plasma. This scheme has been particularly successful on TEXTOR in the so-called Radiative Improved (RI) mode [20]. The reason to prefer low- Z species is that they radiate mainly near the plasma periphery, much more than in the core plasma. This is also why plasma-wetted surfaces are preferably made of a low- Z material.

Impurities are also very useful because they allow the diagnosis of various plasma parameters via spectroscopy, as opposed to the hydrogenic ions, which lose their electron quickly in even a cool plasma.

2.8.2 Wall conditioning

Over the years, several measures have been taken to avoid the detrimental effects of impurities in the plasma. With the advent of better vacuum techniques, the use of low- Z materials for limiters and divertor target plates (particularly carbon), and the improvement of magnetic field geometry, significant progress could be made with respect to impurity control. Started in the mid-1970s, *gettering* with low- Z materials like Li, Be and B is used regularly [21]. The main aim is to lay a film of material on the vessel walls that is more resistant against chemical erosion, while providing a good receptor for oxygen. Oxygen is a persistently stubborn impurity in tokamak plasmas, and originates mainly from water vapor from hidden surfaces not exposed to the plasma or to cleaning discharges. As a result of gettering, oxygen levels in the plasma decrease dramatically (factor 3, to more than 10 on JET), which has been observed both at TEXTOR and JET [21,22]. Metal concentrations may drop with factors of 10 to 20.

In addition, routine baking of the vessel in order to desorb water vapour, and the regular application of a glow discharge in a flow of hydrogen or helium gas can help to decontaminate plasma facing components, especially for oxygen decontamination [22].

The techniques of gettering, baking and glow discharge cleaning, are commonly called *wall conditioning methods*.

2.8.3 Impurity transport

Most impurity ions that are released from the walls quickly return to the surface, but some are transported further into the plasma, and some may even reach the centre of the main plasma. Eventually however, all ions return to the particles sinks. Once the particles enter the plasma, they become ionized, and obey the laws of *impurity transport*, which we now explain in a bit more detail. A distinction has to be made between edge transport and transport in the main plasma. All results mentioned in this section are equally valid for the transport of hydrogenic ions.

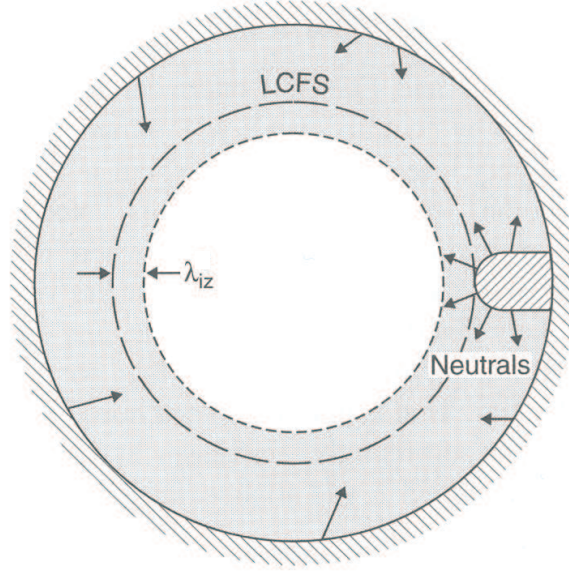


Figure 2.11: The definition of the plasma edge in the context of transport, with indicated ionization depth $\lambda_{i,Z}$.

2.8.3.1 Edge transport

In the context of transport, the plasma ‘edge’ may be defined in two ways.

1. When ionization of neutrals occurs entirely outside the LCFS, then edge transport is considered as the transport that occurs outside the LCFS.
2. If the ionization of neutrals extends inside the LCFS over some distance $\lambda_{i,Z}$, then edge transport is considered to be the transport that occurs outside the radius $a - \lambda_{i,Z}$. $\lambda_{i,Z}$ is the typical ionization depth of an impurity, depending on charge number Z . This is depicted in Figure 2.11.

In the simplest picture, the edge impurity transport can be modelled using the quasi-one-dimensional Engelhardt model, but in general one has to assume a 2D- or even 3D-model. Moreover, in edge transport one has to consider both the transport of ions and neutrals. Edge transport is usually inherently more complex than transport in the main plasma.

We will not go further into edge transport, and rather refer to Ref. [2] for an overview.

2.8.3.2 Main plasma transport

Transport in the main plasma (i.e. the plasma inboard of the edge region) has been extensively studied both theoretically and experimentally. The main plasma transport is usually seen in a 1D cross-field picture, governed by the following equation, containing a diffusion and convection term:

$$\Gamma_{\perp} = -D_{\perp} \frac{dn}{dr} - v_{\text{pinch}} n, \quad (2.2)$$

where Γ_{\perp} is the perpendicular flux density of impurity species with density n , and v_{pinch} is the drift velocity. The transport coefficients D_{\perp} and v_{pinch} can not yet be calculated from first principles, since they are anomalous, and much higher than the neoclassical values. Typically, $D_{\perp} \approx 0.1 - 1 \text{ m}^2/\text{s}$, while $v_{\text{pinch}} \approx 10 \text{ m/s}$ [13]. Hirschman and Sigmar have reviewed the literature on neoclassical transport, see Ref. [23]. An excellent text book series covering classical, neoclassical and anomalous transport has been written by Balescu [24–26].

The transport of helium ions is of particular interest to reactor-grade plasmas, because here the problem rises of an efficient exhaust of the helium ash from the plasma. In order to achieve a stationary and ignited burning D-T plasma, it is necessary that [27, 28]

$$\tau_{\alpha}^*/\tau_E \leq 10,$$

with τ_{α}^* is the global alpha particle confinement time, while τ_E is the energy confinement time. Thus, the particle confinement time should not be ‘too good’. Also, helium neutrals need to be pumped away effectively from some region at the edge.

2.8.4 Measuring impurity content

From the foregoing, it is clear that impurities play a crucial role in tokamak plasmas. In addition, we have seen that theoretical predictions of the impurity distribution throughout the plasma in general can be quite different from the real situation. Therefore, it is very important to obtain reliable measurements of the impurity properties and behaviour in the plasma. Many sophisticated techniques have been developed over the last decades, not only to characterize the impurity behaviour

(transport) itself, but also, using the impurities as a diagnostic, to study particle confinement and deduce such properties as ion temperature and plasma rotation.

A lot of information on impurity transport can be deduced from measurements of the density or concentration of the impurities. Here, concentration refers to the ratio of impurity density to electron density. If absolute concentrations can not be determined, it is often already valuable to have an idea of relative changes in impurity abundances. Relative or absolute impurity concentrations can be obtained using various active or passive spectroscopic techniques [29], such as UV and VUV spectroscopy, bremsstrahlung spectroscopy, Active Beam Spectroscopy, etc. In this work, measurements from bremsstrahlung spectroscopy (Section 3.1) and Charge Exchange Spectroscopy (Section 3.2), which is a form of Active Beam Spectroscopy, are used to assess the impurity content in the plasma.

2.8.5 The ion effective charge Z_{eff}

We now introduce the plasma parameter that is the key quantity of interest in the current work, namely the *ion effective charge* or *average ionic charge* Z_{eff} . Z_{eff} is a local measure for the impurity concentration in the plasma, averaged over all impurities. We will define Z_{eff} on the basis of the electrical plasma resistivity. If an electric field \mathbf{E} is applied in a plasma, collisions between electrons and ions will hinder the acceleration of electrons in response to \mathbf{E} . On the other hand, the plasma ions will, due to their higher mass, remain relatively unresponsive to \mathbf{E} . Thus, just as in a solid, an applied electric field can drive only a limited current, which is reflected in a finite *electrical resistivity* η of the plasma.

We first mention a simple expression for the resistivity of a pure hydrogen plasma, in the presence of an electric field E_{\parallel} parallel to a magnetic field in a cylindrical geometry (or in the absence of a magnetic field), derived from the momentum equation of magnetohydrodynamics³. The parallel resistivity is of the most interest in tokamak plasmas because the plasma current is almost parallel to the magnetic field, while the collisional transfer of momentum between electrons and ions,

³Magnetohydrodynamics, or MHD, is the theory that describes, under certain conditions, the plasma as an electrically conducting fluid. For more information, see a general text book on plasma physics, such as given in Refs. [30] or [31].

due to the presence of an electric field, depends only moderately on the direction of the current.

We assume that the electrons have reached a steady-state situation after application of E_{\parallel} . The resistivity can then be written as [32]

$$\eta_{\parallel} = \frac{m_e \langle \nu_{ei} \rangle}{n_e e^2}, \quad (2.3)$$

where n_e is the electron density and m_e the electron mass. The electron-ion collision frequency ν_{ei} is averaged over an appropriate electron distribution. To calculate the resistivity for an arbitrary orientation of the electric field with respect to the magnetic field, one has to replace the scalar η_{\parallel} by a resistivity with tensorial character. The resistivity perpendicular to the magnetic field is then found to be almost twice η_{\parallel} [33]:

$$\eta_{\perp} \approx 1.96 \eta_{\parallel}.$$

This can be understood because the ions and electrons are much more restricted in their motion across the magnetic field, than along the magnetic field.

Turning again to the parallel resistivity, from the theory of Coulomb collisions, and assuming a Maxwellian electron distribution that is shifted over some drift velocity, one can write (2.3) as

$$\eta_{\parallel} = \frac{\sqrt{2m_e} Z e^2 \ln \Lambda}{12\pi^{3/2} \epsilon_0^2 T_e^{3/2}}; \quad \ln \Lambda \text{ is the Coulomb logarithm,}$$

which is weakly dependent on n_e as well through $\ln \Lambda$. This result overestimates the true resistivity of a hydrogen plasma by a factor of about two. The weakness lies in the assumption of a Maxwellian electron distribution. In reality, the distribution differs somewhat from a Maxwellian because electrons with different velocities respond differently to the combined effect of an electric field and collisions with ions.

In order to obtain a more accurate value for the resistivity, it is necessary to solve the collisional kinetic equation for the electron distribution function, taking into account also electron-electron collisions. The latter do not contribute *directly* to the resistivity, but they rather modify the electron distribution function so as to increase the total drag on electrons due to collisions with ions. Such a (numerical) calculation was done by Spitzer and co-workers [34], who found that for

a pure hydrogen plasma the resistivity is

$$\eta_{\text{H,S},\parallel} \approx \frac{2.8 \times 10^{-8}}{T_e^{3/2}} \Omega\text{m}, \quad T_e \text{ in keV.} \quad (2.4)$$

This is an order of magnitude less than the copper resistivity at temperatures required for thermonuclear fusion.

In practice, several ion species can occur in a tokamak plasma, and we should sum the collision frequency over all ion species. Since the collision frequency for ion species i is proportional to $n_i Z_i^2$, we find that for small impurity concentrations [35]

$$\eta_{\text{S},\parallel} \approx \eta_{\text{H,S},\parallel} \frac{\sum_i n_i Z_i^2}{n_e}. \quad (2.5)$$

The summation starts at $i = 1$ (hydrogenic ions) and goes over all impurity species as well. If we now define the ion effective charge Z_{eff} as

$$Z_{\text{eff}} \equiv \frac{\sum_i n_i Z_i^2}{\sum_i n_i Z_i}, \quad (2.6)$$

and we make use of the quasi-neutrality of the plasma (i.e. $n_e = \sum_i n_i Z_i$), (2.5) becomes simply

$$\eta_{\text{S},\parallel} \approx \eta_{\text{H,S},\parallel} Z_{\text{eff}}.$$

A somewhat more accurate approximation, also valid for higher impurity concentrations, includes a Z_{eff} -dependent factor $N(Z_{\text{eff}})$:

$$\eta_{\text{S}\parallel} \approx \eta_{\text{H,S}\parallel} N(Z_{\text{eff}}) Z_{\text{eff}}.$$

Z_{eff} clearly plays the role of a charge of the plasma averaged over all impurity species. At the same time, it provides an average impurity concentration, weighted by Z_i^2 ⁴. In a pure hydrogen plasma, $Z_{\text{eff}} = 1$, so Z_{eff} can never be smaller than unity. Typical values on present-day machines (e.g. TEXTOR, JET, ASDEX-Upgraded) are $Z_{\text{eff}} \approx 2$.

Returning to the plasma resistivity, we still note that a further modification to the expression (2.5) arises when the toroidal magnetic geometry in a tokamak traps a fraction of the electrons so that they

⁴Note that we can not simply weight by Z_i , since Z_{eff} would become identically 1.

cannot carry a current in response to the electric field [36]. An approximate formula for the resulting neoclassical resistivity is [37]

$$\eta_{\text{neo}} \approx \eta_{\text{H,S}} \frac{Z_{\text{eff}}}{(1-\phi)(1-C\phi)} \frac{1+0.27(Z_{\text{eff}}-1)}{1+0.47(Z_{\text{eff}}-1)},$$

where

$$\begin{aligned} \phi &= \frac{f_{\text{T}}}{1 + (0.58 + 0.20Z_{\text{eff}})\nu_{*e}}, \\ C &= \frac{0.56}{Z_{\text{eff}}} \left(\frac{3.0 - Z_{\text{eff}}}{3.0 + Z_{\text{eff}}} \right), \\ \nu_{*e} &= \epsilon^{-3/2} \frac{R_0 q \nu_{\text{ei}}}{v_{\text{Th,e}}}, \end{aligned}$$

with $v_{\text{Th,e}} = (T_e/m_e)^{1/2}$ the electron thermal velocity and f_{T} the fraction of trapped electrons:

$$f_{\text{T}} = 1 - \frac{(1-\epsilon)^2}{(1-\epsilon^2)^{1/2}(1+1.46\sqrt{\epsilon})}.$$

ϵ is the inverse aspect ratio R_0/a of the tokamak, and q is the so-called *safety factor* [36].

The importance of Z_{eff} in a fusion plasma in general can be summarized as follows. First, there is a direct relation between Z_{eff} and the power radiated from the core plasma through bremsstrahlung (see Equation (3.2)), as well as the dilution of the fuel. As such, Z_{eff} is a crucial quantity for the operation of ITER, the next-step tokamak device (see Section 2.10), and future fusion reactors. For ITER a Z_{eff} value of 1.8 is foreseen. Depending on the discharge scenario, this number may only vary by ± 0.2 , so Z_{eff} should be known with a relatively high accuracy. As far as the fuel dilution is concerned, a consistency check can be performed with the plasma kinetic energy and neutron yield modelling.

Second, Z_{eff} is very important for the study of impurity transport. Even an average Z_{eff} value along a centrally viewing chord is in general weighted heavily by the central plasma. Thus, Z_{eff} generally is a measure for the central impurity content. By measuring impurity fluxes, one can derive transport coefficients (Equation (2.2)).

Finally, since Z_{eff} is linked directly with the plasma resistivity, it also has an important influence on stability and confinement of the plasma.

2.9 Plasma heating

As long as ignition is not reached in a fusion machine, the plasma needs to be heated by means other than the heating by alpha particles. The required heating power is given by [13]

$$P_H = \left(\frac{3nT}{\tau_E} - \frac{1}{4}n^2\langle\sigma v\rangle E_\alpha \right) V,$$

with n the fuel density, T the fuel temperature, $\langle\sigma v\rangle$ the rate of the D-T reaction, E_α the energy carried by the alpha particle per fusion reaction and V the plasma volume. The first term represents the power loss through various mechanisms, the second term is the alpha particle heating.

Several types of heating exist, see Figure 2.12. First, there is the intrinsic ohmic heating by the plasma current because of the plasma resistivity. At low temperatures this heating is very strong, but the resistivity varies with electron temperature as $T_e^{-3/2}$ (Equation (2.4)), so at higher temperatures the effectiveness of ohmic heating quickly drops. Moreover, because the plasma current also plays a role in the confinement, the current density is limited by magnetohydrodynamic stability requirements. Calculation of the ohmic heating power is in practice rather difficult, and depends also on the amount of impurities in the plasma (Z_{eff}). In an ignited plasma, ohmic heating is quite small, amounting to a few MW. To increase plasma energies, on many devices (external) auxiliary heating is applied, up to 25 MW on JET.

A first auxiliary heating method is provided by *Neutral Beam Injection* (NBI). The principle is the injection of neutral hydrogenic particles, which are unaffected by the magnetic field as they enter the plasma. Gradually, the injected atoms become ionized through collisional interaction with the plasma, so that they are forced to follow the magnetic field, and are thus removed from the beam. Through Coulomb collisions, the beam particles transfer their energy to the plasma, thus increasing its temperature. Clearly, as much of the energy carried by the beam as possible should be deposited in the central region of the plasma. Therefore, the absorption of the beam particles should be neither too strong, nor too weak.

Secondly, *Radio Frequency* (RF) *heating* transfers energy to the plasma by electromagnetic waves. When an electromagnetic wave propagates through the plasma, the electric field of the wave accelerates the

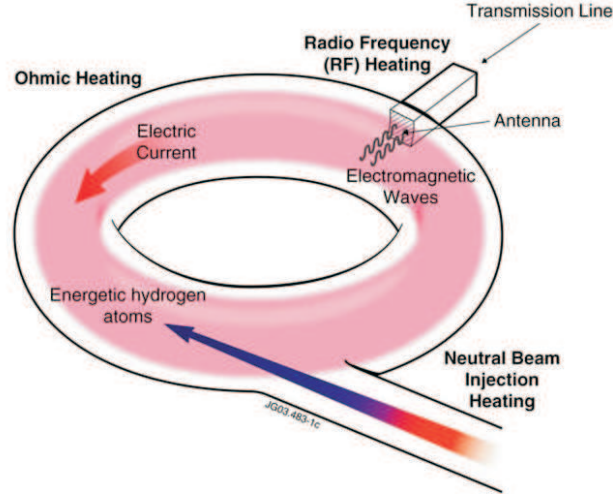


Figure 2.12: Sketch of the various heating methods for a tokamak plasma.

charged particles, which then heat the plasma through collisions. However, just as for ohmic heating, the collisional absorption of electromagnetic waves decreases strongly as a function of plasma temperatures. A solution is provided through the *resonant* absorption of waves, since a magnetized multispecies plasma has a number of resonant frequencies, permitting strong heating of the plasma. Again, power deposition should be the highest in the plasma central region. There exists a large variety of wave modes in a magnetized plasma, so that many different Radio Frequency heating schemes are possible. The most prominent are *Ion Cyclotron Resonance Heating* (ICRH), *Lower Hybrid Resonance Heating* (LHRH) and *Electron Cyclotron Resonance Heating* (ECRH) [38].

2.10 ITER: the next-step tokamak device

A large part of the worldwide efforts on tokamak research is concentrated on the next-step device ITER: International Thermonuclear Experimental Reactor. The aim of ITER is to demonstrate the scientific and technical feasibility of fusion power production. As such, it is meant to be the single step between present-day devices and the DEMO

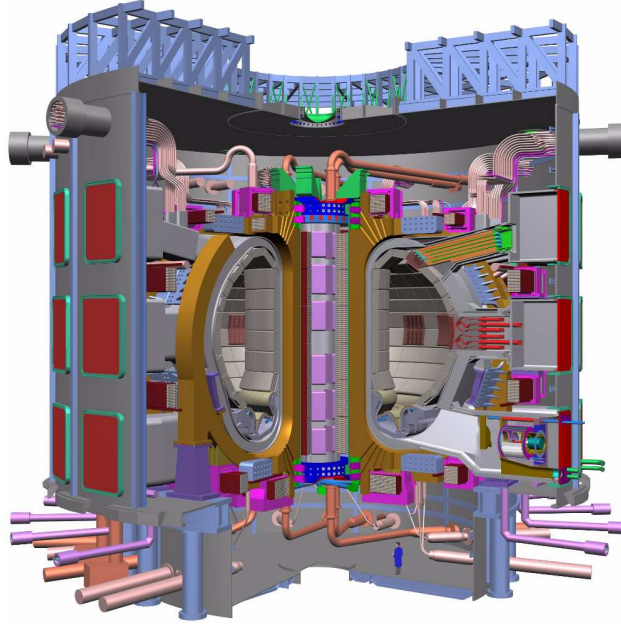


Figure 2.13: ITER cutaway, showing the major components of the tokamak itself.

tokamak reactor concept. Partners in the project are the European Union (EURATOM), Japan, the Peoples Republic of China, India, the Republic of Korea, the Russian Federation and the USA. ITER will be constructed in Europe, at Cadarache in the South of France. The ITER design activities have been officially completed, although several design issues are still a matter of discussion. One of the objectives of ITER is to reach a power amplification factor $Q \geq 10^5$, and demonstrate a steady-state operation with $Q \geq 5$. The actual construction of ITER is planned to commence in the course of 2008, while the first ITER plasma is expected in 2016. Some of ITER's main parameters have been summarized in Table 2.4, while an ITER cutaway is displayed in Figure 2.13.

⁵ Q is defined as the ratio of the produced thermonuclear power to the supplied heating power.

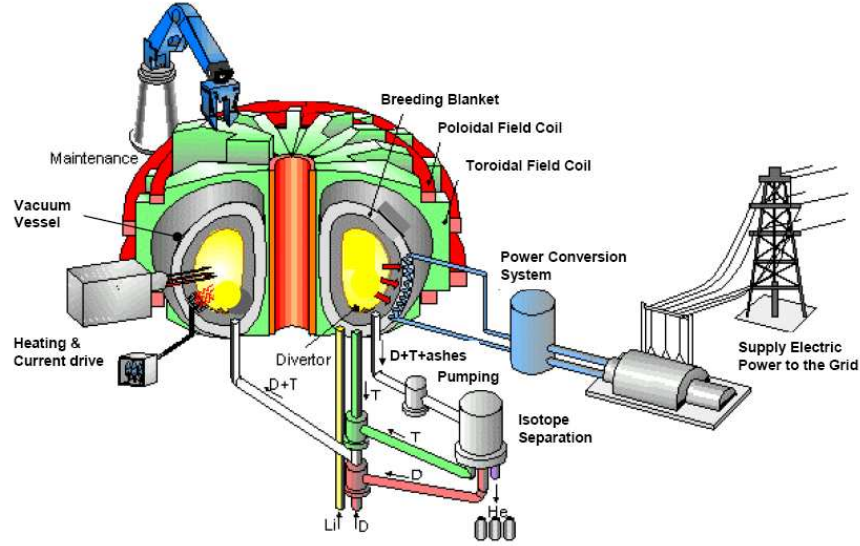


Figure 2.14: A schematic of a tokamak fusion power station.

2.11 Fusion power plants

The eventual main goal of present-day fusion research is to build a power plant, generating electricity from the energy released by fusion reactions in a fusion device [39]. Since the tokamak is currently the most developed concept towards a fusion machine, the first power plants will be based on the tokamak principle. In a fusion power plant, the tokamak is equipped with a divertor. The burning plasma is surrounded by a lithium ‘blanket’, a structure that captures the fast neutrons from the fusion reaction, which, due to their neutrality, escape to tokamak magnetic field. The energy deposited in the blanket by the neutrons will be removed by a heat exchanger, in order to produce electricity in the conventional way. At the same time, the neutrons interact with the blanket to produce tritium, which is also extracted from the blanket and injected, together with deuterium, into the plasma to sustain the fusion process. To minimize dissipation of energy in the toroidal field coils, these coils will be superconducting. The temperature of the plasma will be about 20 keV at a density of approximately 2×10^{20} nuclei/m³. A diagram of a fusion power plant is shown in Figure 2.14.

2.12 The TEXTOR tokamak

Part of the research described in the current work was carried out on the TEXTOR (Torus Experiment for Technology Oriented Research) tokamak at the Institut für Plasmaphysik (IPP) of the Forschungszentrum Jülich (FZJ, Jülich, Germany). TEXTOR operates under the Trilateral Euregio Cluster (TEC), a transnational organizational structure consisting of three EURATOM-associated laboratories, namely the IPP-FZJ, the Laboratory for Plasma Physics of the Royal Military Academy (RMA, Brussels, Belgium) and the Institute for Plasma Physics of the Stichting voor Fundamenteel Onderzoek der Materie (FOM, Nieuwegein, The Netherlands). TEXTOR is a medium-sized limiter tokamak primarily intended for research in the field of plasma-wall interaction. For this purpose, TEXTOR has a number of special design features, such as an excellent access for diagnostics to domains near to the wall, and large access ports suitable for implementing methods to control the plasma boundary. TEXTOR is equipped with auxiliary heating systems (NBI, ICRH, ECRH), a pumped toroidal limiter (ALT⁶-II) and recently the *Dynamic Ergodic Divertor* (DED). The DED consists of a system of sixteen helical coils, installed at the TEXTOR HFS. It generates a perturbation of the magnetic field, resonant to the $q = 3$ surface near the plasma edge. The DED can operate in several different modes ($m/n = 3/1$, $6/2$ and $12/4$), while the perturbation field can be rotated (hence *Dynamic Ergodic Divertor*). The result is an ergodisation of the edge plasma, and a creation of magnetic islands. The main scientific objectives of the DED are:

- study the effect on plasma-wall interaction and confinement in general,
- optimize impurity exhaust,
- optimize the concept of a cold radiating mantle (RI mode),
- distribute the convective heat load on larger areas of plasma-facing surfaces,
- study the effect on particle screening,
- study the effect of different mode structures of the perturbation field.

⁶Advanced Limiter Test

The closed magnetic flux surfaces in TEXTOR are circular.

The plasma-facing surfaces are made of graphite. The by far dominant impurity in many TEXTOR discharges is therefore carbon [40]. Oxygen contents are reduced substantially by gettering methods, see Section 2.8.2. Finally, concentrations of metallic impurities are usually of the order of 0.01%.

Some of the main TEXTOR parameters have been summarized in Table 2.4. An inside view of TEXTOR is shown in Figure 2.15.

2.13 The JET tokamak

The JET (Joint European Torus) device is the largest tokamak in the world, operating under the European Fusion Development Agreement (EFDA) in Culham, UK. The principal aims of the experiment are the investigation of heating and confinement under reactor-relevant plasma conditions, and the study of plasma-wall interaction, alpha particle production, confinement and plasma heating. JET has a pumped divertor configuration and is equipped with NBI, ICRH and LHRH auxiliary heating systems. The closed magnetic flux surfaces are D-shaped, see Figure 2.16. JET is also renowned for its high performance plasmas, particularly during the preliminary tritium experiments, providing 13 MW of peak fusion power.

The limiter surfaces and divertor target plates in JET are fabricated from graphite and CFC. The by far dominant impurity in many JET discharges is therefore carbon [41]. Beryllium is used for gettering, but usually need not be considered as an important impurity. Oxygen contents are reduced significantly by gettering (Section 2.8.2). Finally, concentrations of metallic impurities are generally of the order of 0.01%.

Some of JET's main parameters are listed in Table 2.4. An inside view of JET can be seen in Figure 2.17.

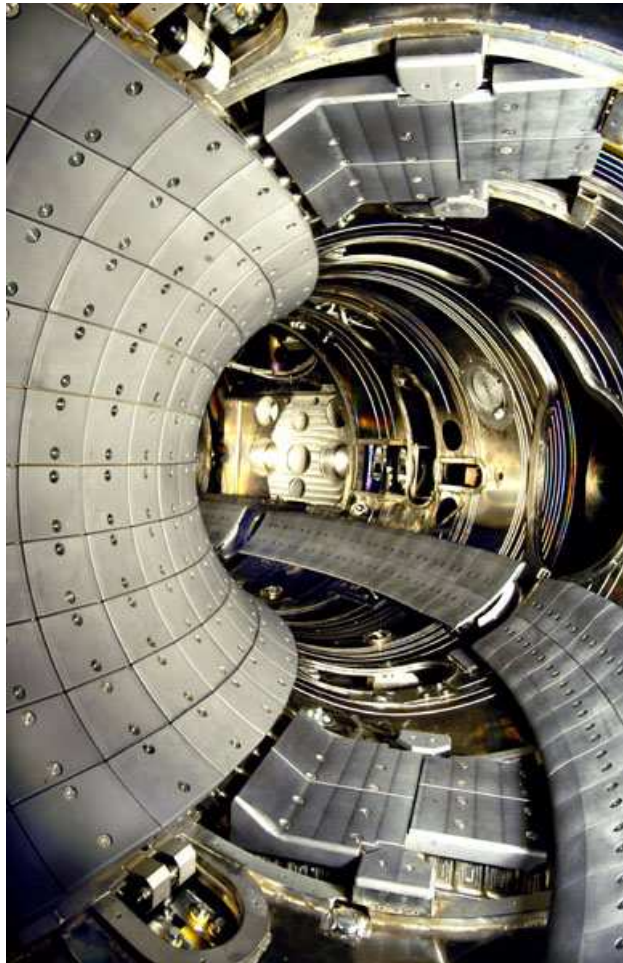


Figure 2.15: Inside view of TEXTOR, with parts of the ALT-II limiter clearly visible on the lower right-hand side, and the graphite tiles covering the DED coils on the left-hand side.

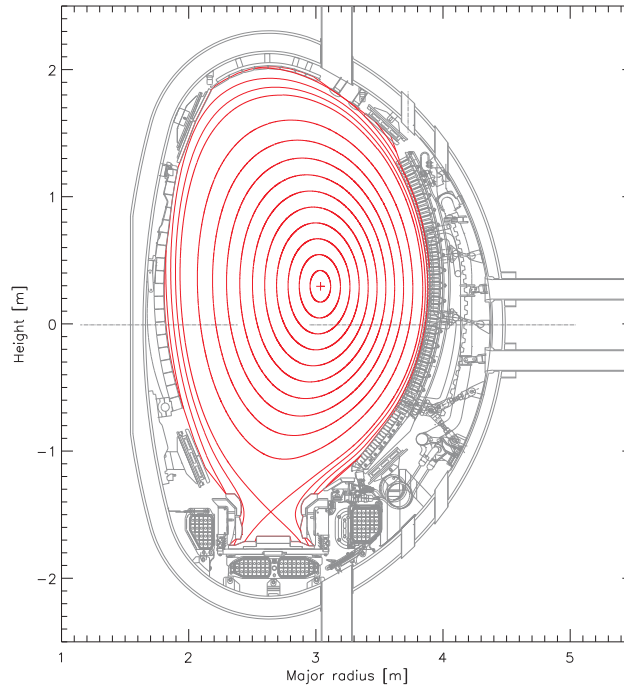


Figure 2.16: Poloidal cross-section of a typical lay-out of the JET magnetic flux surfaces.

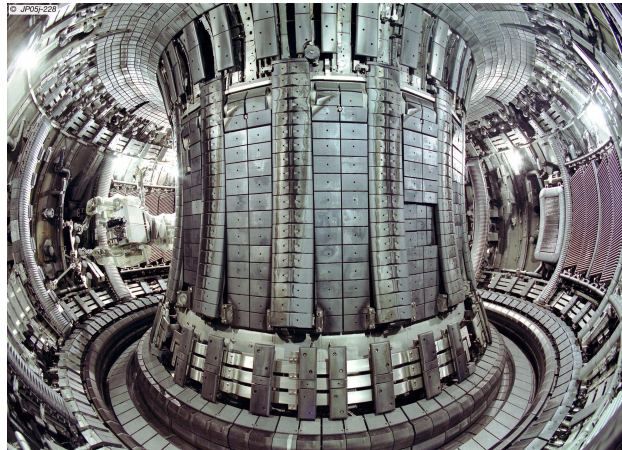


Figure 2.17: Inside view of the JET vacuum vessel, with the divertor chamber clearly visible at the bottom.

Parameter	TEXTOR	JET	ITER
Major radius R_0 (m)	1.75	2.96	6.21
Minor radius a (m)	0.47	1.25	2.00
Plasma volume (m^3)	7	80 – 100	837
Toroidal magnetic field \mathbf{B}_ϕ (T)	3.0	3.4	5.3
Plasma current I_p (MA)	0.8	5 – 7	15 – 17
Pulse duration (s)	< 10	10 – 60	≥ 400

Table 2.4: Typical values of some of the principal parameters for TEXTOR, JET and ITER.

And God said

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}\end{aligned}$$

and there was light!

Maxwell's equations

3

Spectroscopic determination of Z_{eff}

Several plasma diagnostics exist, both passive and active, for the determination of Z_{eff} . In this chapter we concentrate on the two approaches that are most relevant to the current work, namely the deduction of Z_{eff} from visible bremsstrahlung emissivity measurements, and the calculation of Z_{eff} via a summation of impurity densities measured by Charge Exchange Spectroscopy. This chapter will discuss only the measurement principles, while more information on the technical aspects is given in Chapter 4. Other diagnostics for Z_{eff} determination are based on measurements of plasma resistivity, continuum soft-X-rays and neutron yields.

3.1 Visible bremsstrahlung spectroscopy

The most widely used method to assess Z_{eff} is by measuring the bremsstrahlung emissivity from the plasma, usually in the visible part of the spectrum. This technique was first applied by Kadota and co-workers on JIPP T-II in 1980 [42], and soon used worldwide. In the next sections, we will discuss the origin of plasma bremsstrahlung and the actual determination of Z_{eff} from the bremsstrahlung emissivity.

3.1.1 Bremsstrahlung from a plasma

In a fusion plasma, the dominant mechanism responsible for the radiation of bremsstrahlung is the free-free transition of an electron in the electric field of an (impurity) ion. Encounters between two electrons cannot produce radiation by electric or magnetic dipole processes. Instead, the electrons will emit quadrupole radiation, with an associated power that is lower by a factor $(v/c)^2$ —negligible for non-relativistic electrons. An ion, of course, is also accelerated by the field of a passing electron. However, its mass is much greater than the electron's, so its acceleration (and thus its radiated power) is again negligible in comparison to the electron's. We will thus assume that the ions form a stationary background, in which the electrons move.

The path of an electron in the electric field of an ion is hyperbolic. However, on account of the long range of the Coulomb force, it is much more likely for an electron to suffer an appreciable deflection as a result of *many weak interactions* rather than because of a *single close collision*. In practice therefore, a considerable contribution to the total bremsstrahlung emissivity is due to the acceleration caused by distant ions. In a first approximation, the electron trajectory will therefore be a straight line, and this is essentially the Born approximation to the collision problem.

It is possible to derive entirely classically an expression for the bremsstrahlung power from electron-ion collisions. In the non-relativistic limit ($T_e < 10$ keV), this is done for example in Refs. [43] and [44]. Here however, the problem rises of choosing a minimum impact parameter, and a complete quantum mechanical calculation (see e.g. Ref. [44]) results in the following expression for the local brems-

strahlung emissivity¹ ϵ_{ff} due to collisions of electrons with ions charged Z_i :

$$\epsilon_{\text{ff}}(\mathbf{r}, \lambda, Z_i) = n_e(\mathbf{r})n_i(\mathbf{r})Z_i^2 \left(\frac{e^2}{4\pi\epsilon_0} \right)^3 \frac{8\pi}{3\sqrt{3}m_e^2c^2\lambda^2} \left(\frac{2m_e}{\pi T_e(\mathbf{r})} \right)^{1/2} \times e^{-\frac{hc}{\lambda T_e(\mathbf{r})}} \bar{g}_{\text{ff}}[\lambda, Z_i, T_e(\mathbf{r})],$$

with

- $n_e(\mathbf{r})$: electron density,
- $n_i(\mathbf{r})$: ion density,
- $T_e(\mathbf{r})$: electron temperature (in eV),
- $\bar{g}_{\text{ff}}(\lambda, Z_i, T_e)$: the free-free Gaunt factor averaged over a Maxwellian electron velocity distribution at temperature T_e ,
- λ : the observation wavelength.

If we bring the numerical constants together in one factor, we arrive at the following expression for the bremsstrahlung emissivity that is commonly used in plasma physics:

$$\epsilon_{\text{ff}}(\mathbf{r}, \lambda, Z_i) = \frac{1.50 \times 10^{-29} n_e(\mathbf{r})n_i(\mathbf{r})Z_i^2 \bar{g}_{\text{ff}}(\lambda, Z_i, T_e) e^{-\frac{hc}{\lambda T_e(\mathbf{r})}}}{\lambda^2 \sqrt{T_e(\mathbf{r})}} \left(\frac{\text{W}}{\text{cm}^3 \text{sr} \text{\AA}} \right), \quad (3.1)$$

where

- $n_e(\mathbf{r})$: electron density in cm^{-3} ,
- $n_i(\mathbf{r})$: ion density in cm^{-3} ,
- $T_e(\mathbf{r})$: electron temperature in eV,
- λ : observation wavelength in \AA .

¹This term is commonly used in the present context as the power radiated per unit solid angle, per unit of wavelength, per unit volume, but the correct scientific nomenclature, according to C.I.E. standards, is spectral radiant steriscent $L_e^*(\lambda)$ [45]. In the latter scheme, the term ‘emissivity’ denotes the ratio of energy radiated by a blackbody to the theoretical energy predicted by Planck’s law.

The Maxwell-averaged Gaunt factor includes all quantum effects and is, through T_e (and n_e), weakly dependent on the local plasma conditions. It was first calculated by Sommerfeld [46], but the expression is particularly difficult to deal with. Several analytic approximations for \bar{g}_{ff} for hydrogen-like ions are therefore available in different wavelength and temperature regions (see Ref. [47]). In the present case the following formulas are appropriate:

$$\begin{aligned} \text{Griem [48]} &: \bar{g}_{\text{ff}} = \frac{\sqrt{3}}{2\pi} \left[\ln \left(\left[\frac{T_e}{h\frac{c}{\lambda}} \right]^2 \frac{T_e}{Z_i^2 R_y} \right) - \gamma \right] \\ \text{Elwert [47]} &: \bar{g}_{\text{ff}} = \frac{\sqrt{3}}{\pi} \ln \left(\frac{4T_e}{h\frac{c}{\lambda}\xi} \right) \\ \text{Kramer [47]} &: \bar{g}_{\text{ff}} = \frac{\sqrt{3}}{\pi} \left[\frac{3}{2} \ln \left(\frac{T_e}{Z_i^2 R_y} \right) - \ln \left(\frac{h\frac{c}{\lambda}}{4Z_i^2 R_y} \right) - \frac{5}{2}\gamma \right] \\ \text{Ramsey [49]} &: \bar{g}_{\text{ff}} = 3.77 \frac{(0.001T_e)^{0.147}}{Z_i^{0.0579}}, \end{aligned}$$

where

$$\begin{aligned} \gamma &: \text{Euler-Mascheroni constant } (\equiv \ln \xi \approx 0.5772), \\ R_y &: \text{Rydberg energy } \approx 13.6 \text{ eV}. \end{aligned}$$

To determine the most suitable formula, one can compare with the numerical computation by Karzas *et al.* [50]. The wavelength region around $\lambda = 5230 \text{ \AA}$ is usually presumed to be relatively free of line emission (see section 4.1.1.2), and it is therefore especially suited for the determination of Z_{eff} . It turns out (see Ref. [51]) that Elwert's formula results in the best approximation for \bar{g}_{ff} in the case of $Z_i = 1$ in the whole temperature region. For $Z_i = 2$, Ramsey's formula fits the data quite well for temperatures between 300 and 700 eV; for other temperatures Elwert's formula is the most appropriate. For $Z_i = 6$ however (see Figure 3.1), none of the above formulas is a good approximation in the region from 25 to 400 eV. In the region from 500 to 1400 eV Kramer's formula and in the region from 1500 to 2500 eV Griem's formula are the most suitable. Impurity concentrations in tokamaks are in general very low (a few percent or less), and therefore on TEXTOR, for example, Elwert's formula is normally used in the routine calculation of Z_{eff} . It is already clear, however, that the uncertainty in the Gaunt factor is a first issue in the calculation of a reliable value for Z_{eff} .

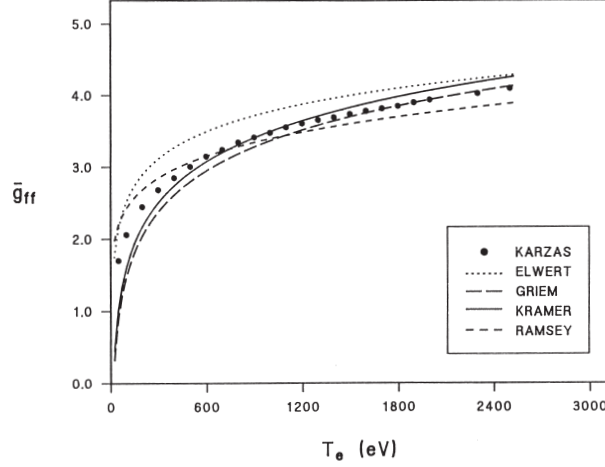


Figure 3.1: The free-free Gaunt factor \bar{g}_{ff} calculated with the approximations of Griem, Elwert, Kramer and Ramsey for $Z_i = 6$ at $\lambda = 5230 \text{ \AA}$. The result of the numerical calculations by Karzas *et al.* is also shown.

3.1.2 Relation with Z_{eff}

In an impure plasma the bremsstrahlung emissivity consists of a sum over all species of contributions of the form (3.1). When the plasma is quasi-neutral, and neglecting the slight Z_i dependence of \bar{g}_{ff} , we can write this as:

$$\epsilon_{\text{ff}}(\mathbf{r}) = \frac{1.50 \times 10^{-29} n_e^2(\mathbf{r}) \bar{g}_{\text{ff}}(\mathbf{r}) e^{-\frac{hc}{\lambda T_e(\mathbf{r})}}}{\lambda^2 \sqrt{T_e(\mathbf{r})}} Z_{\text{eff}}(\mathbf{r}). \quad (3.2)$$

For future purposes, we will gather all constants in a single constant C , and write

$$\epsilon_{\text{ff}} = C \bar{g}_{\text{ff}} \frac{n_e^2 Z_{\text{eff}}}{\sqrt{T_e}}, \quad (3.3)$$

as it can easily be verified that in the full range of possible T_e values, the exponential in (3.2) is unity. The dependence of ϵ_{ff} on T_e is anyhow relatively weak, since the square root in the denominator and the T_e -dependence of \bar{g}_{ff} tend to cancel each other. When also n_e and T_e are known, Z_{eff} can be determined from the bremsstrahlung emissivity ϵ_{ff} . Since Z_{eff} is an absolute number, its determination requires an

absolute calibration of all involved diagnostic systems; the systems for n_e , T_e and ϵ_{ff} measurement.

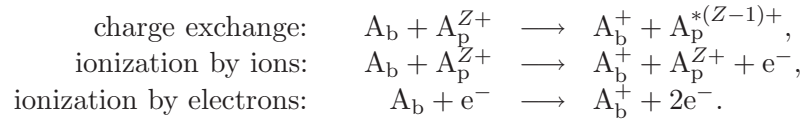
3.2 Charge Exchange Recombination Spectroscopy

In this section we discuss the exciting possibility of determining the impurity content of the plasma by observing the radiation resulting from the interaction of the injected atoms from a neutral beam with the plasma ions. This technique is named *Charge Exchange Recombination Spectroscopy*, and we will demonstrate how a Z_{eff} estimate can in principle be derived from such measurements.

3.2.1 Charge exchange with neutral beam particles

In the interior of hot plasmas, like the plasma found in a tokamak, spectroscopic techniques based on the measurement of atomic line radiation from light impurities, are hampered by the fact that most of the ions are fully stripped. To make inferences about impurity concentrations in the plasma core, one may conduct measurements of line emission measurements from light impurities in the plasma periphery, and derive central abundances via an impurity transport model. However, a more direct way of determining impurity concentrations in the core plasma is provided through interaction of the plasma with a neutral beam of highly energetic hydrogen or deuterium atoms². The neutral beam does not necessarily need to carry a high power, so both a heating beam and a diagnostic neutral beam can be used.

When the hydrogen atoms in a neutral beam enter the plasma, they can be ionized via various atomic processes, and thus be removed from the beam. Let A denote a particle species, while the subscripts b and p refer to beam and plasma particles respectively. The three basic beam ionization mechanisms are then:



²For brevity, we will henceforth denote the beam particles as hydrogen atoms, while keeping in mind that they can also be deuterium atoms.

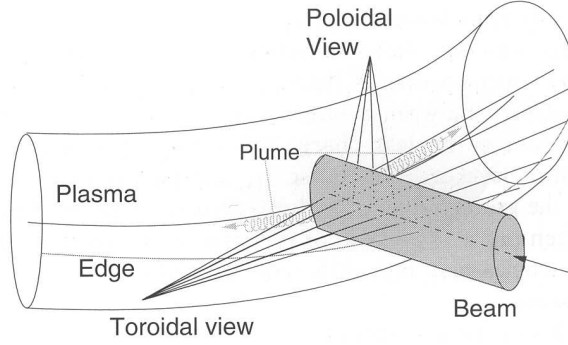


Figure 3.2: Schematic of a possible geometry of lines of sight observing a neutral beam.

We here focus on charge exchange, in which an electron is transferred from the atom to a light plasma ion, such as a carbon, oxygen or boron ion. We will only consider the interaction of the beam particles with fully stripped plasma ions. Both the beam and plasma particle largely retain the energy they had prior to the collision, but the nucleus of the beam particle is now trapped by the magnetic field, and thus leaves the beam. The electron from the beam particle is bound to the plasma particle, which is left in an excited state. The principal quantum number of the recombined state tends to be so high that the emitted radiation during de-excitation is in the visible. This is an important advantage for the diagnostic technique based on the observation of this radiation, namely *Charge Exchange Recombination Spectroscopy*, or CXRS (or shortly CXS). It is a form of active beam spectroscopy.

In practice, the neutral beam is observed at the intersections with a set of lines of sight, as schematized in Figure 3.2. Another advantage of CXS is that the plasma quantities that are derived from it through interaction with the beam, are more or less localized at the intersection of the sight line with the beam volume. Therefore, depending on the beam width and divergence, a relatively localized plasma volume is observed, so that radial profiles can directly be calculated without the need for an inversion method.

3.2.2 Calculation of light impurity densities

Several plasma quantities can be derived from the observation of the charge exchange radiation, which are essentially the impurity temperature and flow velocity, as well as the impurity density. The determination of local absolute impurity densities (or concentrations) is particularly difficult.

The first difficulty is that one needs to distinguish between the charge exchange emission coming from the interaction volume of the beam with the plasma, and various other sources of radiation in the same wavelength region. The latter includes the bremsstrahlung background from the entire line of sight, atomic lines from other impurity species and lines resulting from collisional excitation of the impurity of interest. Also the so-called *plume effect* needs to be taken into account, which originates as follows. A plasma impurity that has undergone a charge exchange reaction with a beam particle decays quickly to its ground state. However, even before it gets reionized, it can, being a charged particle, follow the magnetic field lines to a region outside the beam volume. There, it can radiate again through collisional excitation. The plume effect (indicated in Figure 3.2) is most difficult to discriminate. On the other hand, there is also some line radiation observed near the LCFS from charge exchange reactions between fully stripped ions and a population of neutrals in excited states. This is the source of the *passive* component of the CX spectrum (i.e. not beam-related), which also has to be discriminated from the active part. We will for the most part of this work consider only the contribution of fully stripped carbon to Z_{eff} , and Figure 3.3 shows a typical CX spectrum of CVI in a JET discharge. Incidentally, the bremsstrahlung background from a CXS diagnostic provides also a possibility for the determination of Z_{eff} profiles, just like a dedicated diagnostic for bremsstrahlung measurement. In Chapter 6, we will even follow precisely this route for the determination of Z_{eff} from continuum radiation.

Secondly, one needs to know the absolute beam intensity along the beam. The beam is generally substantially attenuated by the plasma. The beam attenuation depends on the cross-section for the various beam ionization processes, giving rise to a beam stopping cross-section $\sigma_{s,Z}$ for an ion species charged Z . The local beam density n_b is related

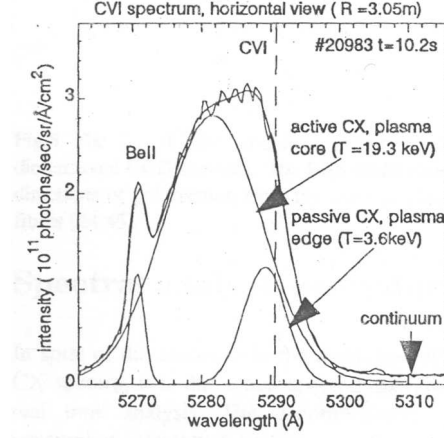


Figure 3.3: Charge exchange spectrum of CVI at 5290.5Å ($T_e = 19.3$ keV) from the horizontal JET CX diagnostic (see Section 4.2). Note the active (beam-related) and passive contributions to the spectrum ($T_e = 3.6$ keV) [3].

to the vacuum density $n_{b,0}$ by

$$n_b = n_{b,0} \exp \left(- \int n_e(s) \left[\sigma_{s,e} + \sum_i \sigma_{s,Z_i} \frac{n_i(s)}{n_e(s)} \right] ds \right), \quad (3.4)$$

where s denotes a length parameters along the neutral beam path, and n_i is the density of ion species i . $\sigma_{s,e}$ denotes the electron-induced beam stopping cross-section. Clearly, a good knowledge of the electron density along the beam is crucial for the calculation of the beam attenuation. In addition, the species mix of the beam and its spatial profile are quantities that need to be measured and may depend on the beam operating conditions.

Finally, the effective emission rate $\langle \sigma v \rangle_{CX}$ for the CX line under study has to be assessed [41]. The emission rate must take into account the cross section of capture into the upper quantum level of interest, as well as the branching ratio of the subsequent radiative deexcitation. In addition, so-called cascade processes may be important, in which a charge exchange reaction places an electron on a high quantum level, which subsequently populates the upper level of the line of interest. However, there are practically no independent measurements of the charge exchange emission rates, so one has to rely on theory, see e.g.

Refs. [29] and [52].

The lines of sight of a CX diagnostic are directed such that they are more or less tangential to the magnetic flux surfaces at the crossing with the beam. In this case, the observed CX spectral intensity Φ_{CX} for the impurity species charged Z is related to the impurity density n_Z by

$$\Phi_{\text{CX}}(\lambda)\Delta\lambda = \frac{1}{4\pi}n_Z\frac{\Delta\lambda}{\sqrt{\pi}\lambda_D}\exp\left[-\frac{(\lambda-\lambda_0)^2}{\lambda_D^2}\sum_E\langle\sigma v(E)\rangle_{\text{CX}}\right. \\ \left.\times\int n_b(E,s)ds\right],$$

where λ is the observation wavelength, λ_D is the Doppler width of the spectral line and λ_0 is the wavelength at the line peak. The integral is over the intersection of the line of sight with the beam profile. A summation has to be performed over the beam energy E because a neutral beam generally consists of several energy components.

As noted in (3.4), the local beam density depends on the stopping cross-sections and densities of *all impurities*. Therefore, the impurity densities are determined from the measured absolute spectral intensities and calculated neutral beam densities, by an iterative procedure [3, 41]. An initial estimate for the impurity levels is usually obtained from visible bremsstrahlung measurements. The calculations can be performed using the Charge Exchange Analysis Package (CHEAP), a code originally written at JET.

3.2.3 Calculation of Z_{eff}

The definition of Z_{eff} , Equation (2.6), can also be rewritten as

$$Z_{\text{eff}} = 1 + \sum_i Z_i(Z_i - 1)\frac{n_i}{n_e}, \quad (3.5)$$

where the sum is over all impurity species. If we take only fully stripped carbon into account, we get

$$Z_{\text{eff}} = 1 + Z_C(Z_C - 1)\frac{n_C}{n_e} = 1 + 30\frac{n_C}{n_e}.$$

This will be of later use.

3.2.4 Other parameters derived from CXS

Charge Exchange Spectroscopy is also commonly used for the measurement of impurity temperatures and flows. These depend on the Doppler width and position of the spectral lines, respectively. More information can be found in Refs. [41], [53] and [54].

3.2.5 A note on data availability

The analysis of CX spectra for the calculation of core impurity concentrations has first been performed on JET [41], and a rich database of processed quantities from CX measurements is available. However, the difficulties in the determination of impurity densities start already with the calibration of the diagnostic system. Uncertainties on the calibration are a first important source of eventual uncertainties on the impurity concentrations. The JET CX systems have been calibrated with great care on several occasions. On the other hand, during the last years no such reliable calibration has been performed for the TEXTOR CX system³.

Moreover, the diagnostic for bremsstrahlung measurement in the visible, described in Section 4.1.1.2, is new on TEXTOR. As we will discuss, there have been several initial problems (reflections, calibration), leading to erroneous line-integrated bremsstrahlung emissivities, thus preventing the reconstruction of Z_{eff} profiles. Only in a later stage of the present work were these issues solved.

Therefore, the study of the integrated estimation of Z_{eff} from measurements of bremsstrahlung and CX impurity densities (Chapter 6) has been performed using data from the JET CX diagnostic. A brief overview of this system is given in Section 4.2.

Since on JET carbon is by far the dominant impurity (see Section 2.13), we will in most of the remainder of this work consider only the contribution of carbon to Z_{eff} . Moreover, at the typical temperatures in a large part of the plasma cross-section at JET, we need to take into account only fully stripped carbon, and we can neglect the other charge states. In the JET discharges that we will consider, only fully stripped carbon was monitored by the CX system. If also other

³This situation will change with the operation of the new system for active beam spectroscopy that has been built on TEXTOR during spring 2006. This system is a pilot experiment for the diagnostic for active beam spectroscopy that will be installed on ITER. The commissioning is expected to start in fall 2006.

impurities give an appreciable contribution to Z_{eff} , their influence can still be modelled under some circumstances, see Section 6.4.3.

3.3 Inconsistency between the continuum Z_{eff} and the CX Z_{eff}

The measurement of bremsstrahlung emissivity and of core impurity concentrations via CXs are the two most popular ways for the determination of Z_{eff} . Since the two methods rely on a different measurement principle, many of the sources of error on Z_{eff} are of a very different nature. Unfortunately, this leads in most cases to an inconsistency between the Z_{eff} value assessed from bremsstrahlung measurements (which we will call the *continuum* Z_{eff}) and the Z_{eff} calculated from CX impurity density measurements (which we will refer to as the *CX* Z_{eff}).

The inconsistency between the continuum Z_{eff} and the CX Z_{eff} appears to be a general problem, at various machines [55–57], and we will give here an example from the JET tokamak. Figure 3.4 shows, as a function of time, the Z_{eff} from visible bremsstrahlung, line averaged along a centrally looking chord, and its equivalent CX line integral, reconstructed from contributions of C^{6+} , Be^{4+} and He^{2+} . As in most cases, the CX Z_{eff} lies systematically below the continuum Z_{eff} . The presence of other impurities is in some cases the cause of a considerable underestimation of the CX Z_{eff} (see below). Apart from this quantitative discrepancy between the two Z_{eff} signals, there is even a certain qualitative difference. Indeed, whereas the continuum Z_{eff} has a continuously rising tendency, the CX Z_{eff} remains basically constant after $t = 14\text{s}$.

Another example from JET is shown in Figure 3.5. For this discharge, the continuum and CX Z_{eff} time traces are shown on the magnetic axis. The continuum Z_{eff} has been calculated from the inversion of the bremsstrahlung line-integrals obtained from the baseline level of the CX spectra. In this discharge, only fully stripped carbon (C^{6+}) was taken into account as an impurity. Clearly, in this particular instance, the temporal evolution of the continuum Z_{eff} and the CX Z_{eff} is completely different.

A last example is shown in Figure 3.6. Here, the continuum Z_{eff} on the magnetic axis is on the average a factor of 2.5 higher than the CX Z_{eff} . Since in this instance the continuum Z_{eff} is generally rather high,

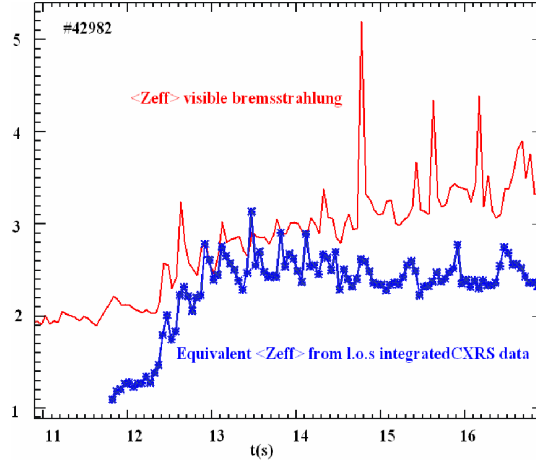


Figure 3.4: Line-averaged continuum Z_{eff} and equivalent CX Z_{eff} line integral for JET #42982.

reaching peak values of over 9, it is clear that the Z_{eff} discrepancy here is at least partly due to a considerable overestimation of the continuum Z_{eff} . This will be confirmed by the analysis in Chapter 6.

The reason for the observed inconsistency of Z_{eff} values lies in the various uncertainties that enter the derivation of Z_{eff} , for both methods. For the determination of Z_{eff} from bremsstrahlung measurements, the following sources of uncertainty can be identified, listed here in descending order of estimated importance.

- It follows from the discussion in Section 4.1.3 that the reconstruction of emissivity profiles can be a large source of error. This is also true for central emissivity values in the profile, due to the propagation of errors from the edge. In addition, there may be poloidal asymmetries in the plasma, so that the emissivity can no longer be assumed to be constant on magnetic flux surfaces. This is of special relevance in the case of DED ergodized edge plasmas on TEXTOR. Moreover, the reconstruction critically rests on the knowledge of the magnetic equilibrium, which, it too, always depends on how accurate the true equilibrium was modelled.
- Uncertainty on the n_e and T_e profiles that are used for the calculation of the Z_{eff} profile, introduce further errors. Due to the quadratic dependence of ϵ_{ff} on n_e , especially the n_e profile should

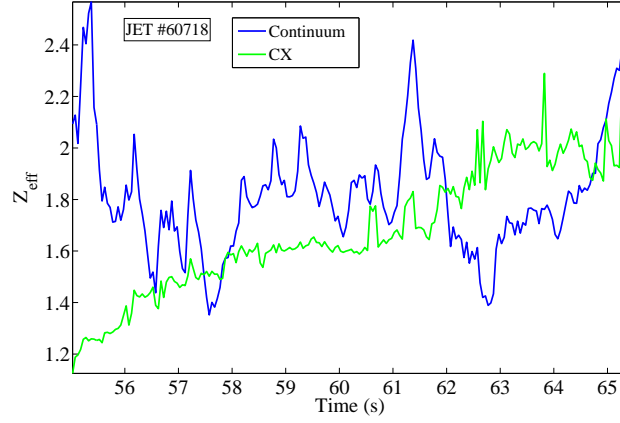


Figure 3.5: On-axis continuum Z_{eff} and CX Z_{eff} in a carbon-dominated plasma for JET #60718.

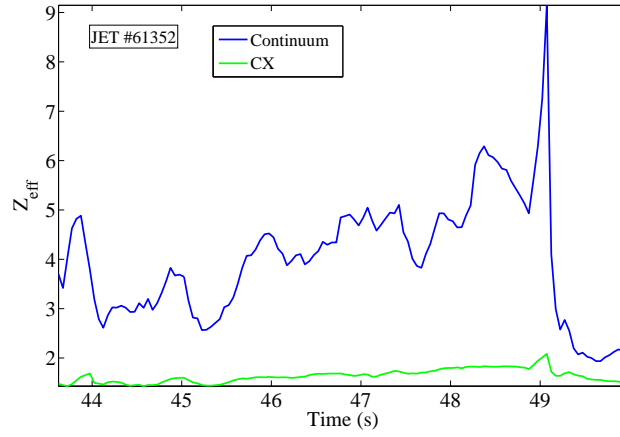


Figure 3.6: On-axis continuum Z_{eff} and CX Z_{eff} in a carbon-dominated plasma for JET #61352.

be known relatively accurately. This can be an issue, particularly near the plasma boundary. The difference in toroidal location of the diagnostic for the measurement of ϵ_{ff} , n_e and T_e can cause even more inaccuracies.

- The spectral window that is used for the determination of the bremsstrahlung continuum should be free of atomic line radiation. On TEXTOR, for example, this is verified regularly, but it is still possible that under certain plasma conditions, this requirement is no longer fulfilled. If the bremsstrahlung emissivity is determined from the background of the CX spectrum, then errors are introduced because one has to separate the background from the rest of the CX spectrum. Furthermore, we will see in Section 4.1.3 that the contribution of non-bremsstrahlung edge components in the continuum can greatly influence the reconstructed profile, an influence that can propagate also towards the centre of the profile. Examples of these components are recombination radiation, molecular bands and black-body radiation from hot material components inside the tokamak vessel.
- As described in Section 4.1.2, the relative and absolute calibration of the bremsstrahlung system also represent a source of errors.
- There can be several issues related to a suboptimal design of the involved diagnostic hardware.
- A long-standing issue with the measurement of bremsstrahlung in the visible is the reflection of the plasma light on various surfaces inside the tokamak vessel. Reflections can be minimized by mounting a viewing dump on the wall that is seen by the diagnostic, as demonstrated in Section 4.1.1.2, but it is impossible to eliminate all reflections physically. By assuming a reflection model for the bremsstrahlung on the vessel surfaces, it is however possible to eliminate most of the influence of reflections on a derived Z_{eff} profile, see Ref. [58]. On TEXTOR this does not lead to much improvement as the installed viewing dump already eliminates most of the reflections, see Section 4.1.1.2.
- As mentioned in Section 3.1.1, in the derivation of the bremsstrahlung emissivity an approximation has to be chosen for the

Gaunt factor. This represents an uncertainty in the physical model, although it is estimated that the effect is modest.

- The electronics that are used in the experimental setup introduce a certain amount of measurement noise. This is estimated to be of minor importance.

As far as the calculation of the CX Z_{eff} is concerned, the following list gives the uncertainties that can be introduced, again in descending order of possible impact.

- One of the main issues is the determination of a correct fit to the active CX spectral lines that are used to monitor the impurity species of interest. This subject was already touched upon in Section 3.2.2.
- Also here, the analysis depends on the knowledge of the magnetic equilibrium, in order to assess the geometry of the lines of sight with respect to the magnetic flux surfaces.
- Although in many TEXTOR plasmas and particularly in JET plasmas, fully stripped carbon is the main impurity that gives by far the dominant contribution to Z_{eff} , the influence of other impurity species might not be negligible. If a certain impurity species, occurring in the plasma with a significant abundance, is not monitored by the CX system, then this will result in an underestimation of Z_{eff} .
- The n_e profile enters the derivation of the CX Z_{eff} through calculation of the neutral beam attenuation, see (3.4). Due to the exponential dependence, uncertainties in the n_e profile will introduce uncertainty in the beam attenuation and the calculated Z_{eff} . The dependence of the CX Z_{eff} on n_e is more-or-less linear, so the influence of n_e uncertainties on the CX Z_{eff} is in any case less important than for the continuum Z_{eff} . The influence of T_e profiles is not very important.
- The relative channel-to-channel calibration, as well as the absolute calibration are a further source of error.
- As is the case for the bremsstrahlung system, there may be important issues related to the design of the CX diagnostic hardware.

- Uncertainty is introduced by the measurement noise.
- The derivation of the CX Z_{eff} rests on the determination of atomic data (cross-sections, rate coefficients). It is however estimated that these atomic data are accurate enough to introduce only minor uncertainty in the eventual CX Z_{eff} .

Thus, both the continuum Z_{eff} and the CX Z_{eff} are influenced by many sources of both statistical and systematic uncertainty. It is cautious not to put too much trust in either of the two Z_{eff} estimates. We therefore conclude this section with the following statements.

The inconsistency of the continuum Z_{eff} and the CX Z_{eff} is a long-standing problem. The determination of a consistent Z_{eff} profile that is reliable over the whole plasma cross-section, is at present a real challenge.

The purpose of the work described in Chapter 6, is to tackle some of the issues related to Z_{eff} determination using recent advanced methods from statistical data analysis.

If your experiment needs statistics, you ought to have done a better experiment.

Ernest Rutherford

4

Diagnostic systems

In this chapter, an overview is given of the diagnostic for visible bremsstrahlung measurement on TEXTOR, which was developed in the context of this work. A short paragraph on the CXS diagnostic on JET is also included. The data that are used in Chapter 6 for the integrated estimation of Z_{eff} , have been recorded using this diagnostic.

4.1 The visible bremsstrahlung diagnostic on TEXTOR

We now proceed with an overview of the main aspects of the design, construction and operation of a diagnostic for bremsstrahlung emissivity measurements in the visible on TEXTOR. The preparation and commissioning of this diagnostic was carried out entirely within the frame of the present work, and so we will conduct the discussion in some detail. The diagnostic is in routine operation on TEXTOR, running entirely automatically, apart from the occasional maintenance work. It routinely yields a set of line-integrated bremsstrahlung emissivities, available on-line via the TEC Web Umbrella (see below), as well as a time trace for a line-averaged Z_{eff} along a single centrally viewing chord. In addition, from the set of emissivity line-integrals a radial emissivity profile can be reconstructed. This is usually done using an Abel inversion routine, but has recently also been performed by Tikhonov and Maximum Entropy regularized inversion. Using also profiles for the electron density and electron temperature, a radial Z_{eff} profile can finally be calculated. The current bremsstrahlung Z_{eff} diagnostic is described in Refs. [59] and [60].

4.1.1 Experimental set-up

4.1.1.1 The pre-TEXTOR-DED diagnostic

The diagnostic for bremsstrahlung emissivity measurement that has been in routine operation until the installation of the Dynamic Ergodic Divertor at TEXTOR, and which is the predecessor of the current diagnostic, is described in detail in [51]. This diagnostic consisted of two parts, the first comprising a set of seven fixed chords, the second featuring an oscillating mirror scanning a poloidal plasma cross-section. The light was in both cases transported to a diagnostic room outside the TEXTOR bunker, a small wavelength band in the visible was selected through a Fabry-Pérot interference filter, and the light was detected using photomultiplier tubes.

With the installation of DED, a number of issues arose with respect to the design of this diagnostic, among which we mention here the main three:

- the typical repetition time of the mirror diagnostic was only 150

ms, which makes it difficult to study plasma phenomena that are subject to rapid variations.

- the system with seven fixed chords had a much higher time resolution (typically 50 μ s), but using only seven chords it can become difficult to reconstruct a sensible Z_{eff} profile. If one increases the number of line-integrals available for the Abel inversion routine (see below), the accuracy of the reconstruction of Z_{eff} profiles can be improved.
- for reasons related to the DED installation itself, it was required that the diagnostic be moved from toroidal section 7–8 to section 12–13. This location is very near one of the DED coil feed-throughs, where operation of the stepping motor driving the oscillating mirror would have been hampered by magnetic forces. The design with the oscillating mirror thus had to be abandoned.

These issues influenced the main aspects of the design of an upgraded diagnostic, as described in the following section. Eventually, we chose a design where the time resolution could be increased with respect to the time resolution of the mirror diagnostic, while at the same time the number of chords could be enhanced compared to the previous seven fixed chords.

4.1.1.2 The upgraded visible bremsstrahlung diagnostic

In order to solve the issues associated with the previous diagnostic, it was decided to install a number (24 in effect) of fixed channels looking into a single poloidal plane, with the detection being made by a CCD camera. This way, the number of channels can be maximized while keeping the total cost of the diagnostic within range, as opposed to the acquisition using relatively expensive photomultiplier tubes. At the same time, the disadvantages of the oscillating mirror pose no longer an issue. In addition, a few channels throughout the observed poloidal cross-section have been coupled to photomultipliers, in order to allow the detection of transient phenomena.

We next discuss the different parts of the diagnostic, including the necessary software, starting from the TEXTOR side. A schematic overview of the diagnostic is shown in Figure 4.1.

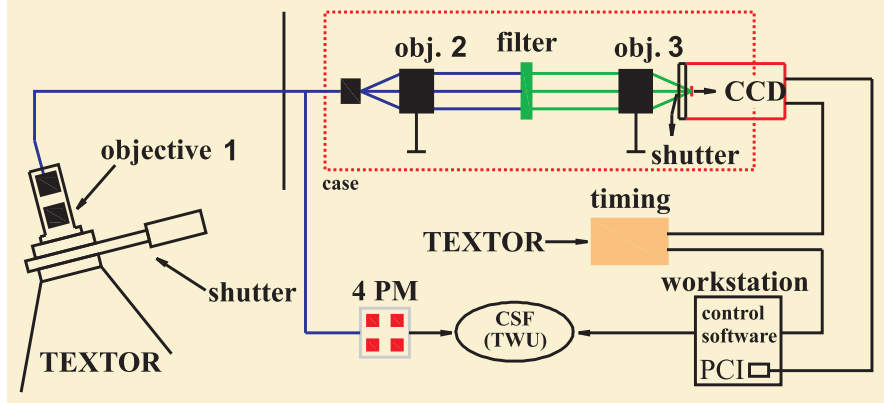


Figure 4.1: Overview of the main components of the visible bremsstrahlung Z_{eff} diagnostic on TEXTOR.

TEXTOR side The light from the plasma is focussed on an array of 38 fibre optic cables (CeramOptec Optran PUV, core diameter 600 μm), located in section 12–13 of the tokamak, and on top of the vacuum vessel (see Figure 4.3). The focussing is done by means of a commercial Nikon AF Nikkor 85 f/1.8D objective, focussed on the equatorial plane of the torus. The support for the fibre holder and the objective can be rotated about an axis in order to be able to slightly incline the plane defined by the chords with respect to a poloidal cross-section. This permits the positioning of the chords in the middle of the 15 cm gap between two of the ALT-II toroidal limiter blades, so as to minimize the interference in any channel (by line radiation, recombination radiation, etc.) originating from the plasma region in the vicinity of ALT¹. In addition, the support can be rotated about a second axis, perpendicular to the one mentioned above, which allows moving the chords within the observational plane. This support was upgraded in the summer of 2005, and the final geometric configuration of the lines of sight is shown in Figure 4.4. Once a common observation point has been chosen for all lines of sight, a chord can be identified by its *impact factor*, which is the distance between the vessel centre, and the intersection of the chord normal, which passes through the vessel centre. An alternative way to label a chord is by measuring its intercept on the equatorial

¹This interfering radiation posed some problems for the operation of the previous diagnostic, as described in [51].

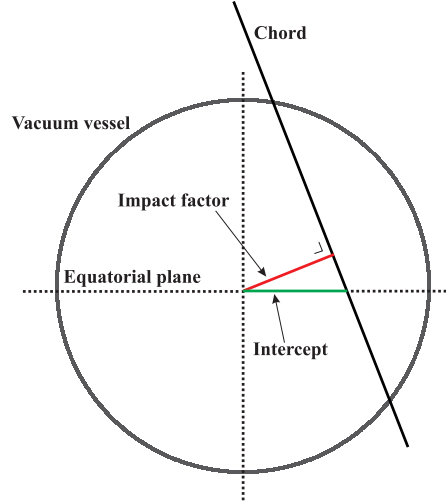


Figure 4.2: Impact factor and intercept (relative to the the centre of the vacuum vessel) for an example chord viewing a plasma cross-section.

plane, relative to the vessel centre. The intercept can also be measured with respect to the centre of the plasma column. Both the impact factor and intercept of an example chord are indicated in Figure 4.2. Originally, 28 channels were intended to be coupled to a CCD camera, and ten to photomultiplier tubes, but the current configuration permits only 24 channels, focussed on the CCD, to be illuminated, as well as seven more channels for use with photomultipliers. This gives a total of 31 channels that can be illuminated at present. The other channels are at the moment cut off by the liner surface. Further improvements to the optical set-up at the TEXTOR side can augment the number of illuminated channels, in order to make use of the full fibre array. The access port where the diagnostic is situated, is equipped with a shutter, which closes automatically during wall conditioning by glow discharge. This prevents the deposition on the window of material sputtered from the electrodes that sustain the discharge, aiding in retaining a nearly constant window transmittance. Nevertheless, the window transmittance also decreases during normal tokamak operation. This effect is however small, and up to now there has been no need to remove the window for cleaning.

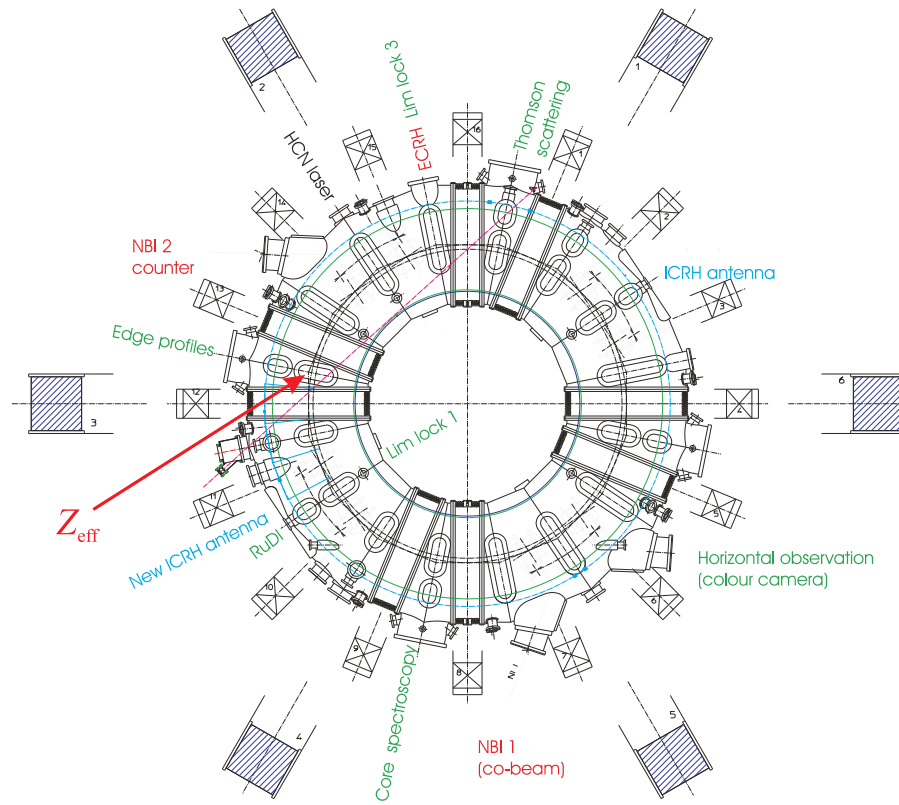


Figure 4.3: Top view of the TEXTOR vessel, with the port for the visible bremsstrahlung Z_{eff} diagnostic indicated ($'Z_{\text{eff}}'$).

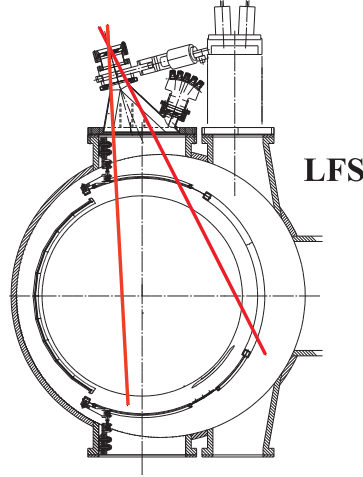


Figure 4.4: The viewing geometry of the TEXTOR visible bremsstrahlung diagnostic in a poloidal cross-section. The red lines indicate the field of view containing the 31 lines of sight.

Camera side: optics The light that enters the fibres is transported over a distance of approximately ten meters to a diagnostic room outside the TEXTOR bunker. The fibres on the camera side are arranged in a rectangular packing, and the exit of the fibres lie in the focal plane of a second Nikon objective (AF Nikkor 85 f/1.8D). The light then passes, in a parallel bundle, a highly blocked interference filter (Barr Associates Inc.) with a peak transmission of 67 percent at a wavelength of 5230.1 \AA , and a FWHM of 21.9 \AA . The filter selects a narrow wavelength band that is known to be relatively free of line emission on TEXTOR. The reason to create a parallel bundle is that the transmission of the filter depends on incidence angle. The peak transmission is at 5230.1 \AA only for normal incidence. The wavelength at which the filter reaches its peak transmission, will henceforth be designated by λ_{ff} . The wavelength range around λ_{ff} is checked at regular intervals to ensure that no atomic lines interfere with the continuum. This can be seen on a typical TEXTOR spectrum, shown in Figure 4.5. A final objective (Nikon AF Nikkor 50 f/1.4D) then focuses the light from 24 channels on a CCD camera. Four more channels are for the moment connected to photomultiplier tubes, with the direct possibility to raise this number to seven.

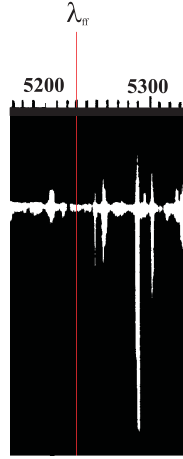


Figure 4.5: Part of the TEXTOR line spectrum, indicating no spectral lines in the immediate vicinity of λ_{ff} . The wavelengths are indicated in Å.

Both objectives at the camera side are mounted on horizontally moveable tables equipped with micrometer screws so as to allow an accurate focussing. A high focussing accuracy is important:

1. for objective 2 in Figure 4.1 to ensure that the light passes through the interference filter in a parallel bundle,
2. for objective 3 in Figure 4.1 to ensure a sharp image of the fibre exits on the CCD, to minimize channel cross-talk at this point.

The entire system at the camera side is enclosed in a PVC case to prevent ambient stray light from entering.

CCD camera

The principle of CCD imaging A *Charge-Coupled Device* (CCD) (see e.g. Ref. [61]) is a solid state image sensor with a very high sensitivity, making it an ideal detector in situations where illumination levels are low. The line-integrated bremsstrahlung emissivities on TEXTOR are indeed low; of the order of $10^{-9} \text{ W}/(\text{cm}^2 \text{ sr } \text{\AA})$. In addition, CCD sensors provide a very good linearity.

A CCD chip is manufactured from a thin slice of very pure silicon. The silicon surface is divided in many light-sensitive areas (the pixels) in a rectangular checkerboard fashion, by laying down narrow strips of dopants that render the silicon photosensitive. Columns of pixels in a so-called *full-frame* sensor are called *vertical registers*. Photons striking the sensor can transfer electrons in the conduction band of the silicon, via the photoelectric effect. If a positive voltage is applied to the strips, potential wells are created, trapping the free electrons. As photons fall onto the chip, an image builds up in the silicon. In order to read out the image, the voltages applied to the strips are changed so that the electrons move from one potential well to the other, a process called *clocking*. At the top (or bottom) of the CCD, a special row of potential wells is constructed, the *horizontal register*. The charge is thus moved up the vertical register, line by line into the horizontal register, which is subsequently read out pixel by pixel. This process is schematized in Figure 4.6. The read-out charge is fed to an amplifier and digitized by an Analog-to-Digital Converter (ADC) for subsequent processing by a computer.

The DTA iCAMII camera The camera used in our application is a full-frame iCAMII CCD camera by DTA s.r.l., Italy, equipped with a front-illuminated Kodak KAF400E chip and a 12 bit ADC. The CCD, consisting of 768×512 pixels, can be cooled by a single-stage Peltier cooling element (40°C below ambient temperature) in order to increase the signal-to-noise ratio. The iCAMII camera has a very low dark current of 0.6 electrons per pixel per second at -5°C . In order to obtain a maximum signal-to-noise ratio, the pixels are binned in groups of 8×8 . The CCD is usually kept at 0°C , where the typical signal-to-noise ratio for 50 ms exposure time and 8×8 binning at a read-out speed of 4.0 Mpixel/s, is 45 dB. Thus, the camera is sufficiently sensitive to yield a good signal-to-noise ratio, although the bremsstrahlung

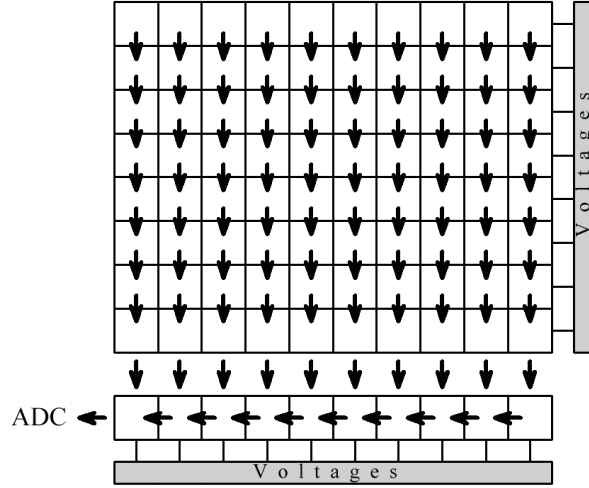


Figure 4.6: The principle of imaging and charge read-out in a CCD sensor, with the horizontal register at the bottom of the array.

emissivity yield at the CCD camera is relatively low. Partly, this is due to the low bremsstrahlung emissivity of the plasma itself. On the other hand, there can be a lot of losses in the objectives (containing many lenses) and in the light guides. When the system is idle, the CCD chip is protected from dust and any residual stray light by an external electromechanical shutter, which opens only during acquisition. The main features of the camera are summarized in Table 4.1.

The computer and control software The camera is connected via a 16 bit PCI² interface to a Hewlett-Packard XW6000 workstation, with two Intel Xeon processors running at 2.4 GHz. It works under the Microsoft Windows 2000 operating system. The routines for controlling the camera and processing the images run entirely on one of these processors, so that the other can be used for running other applications, while keeping full processing power for running the diagnostic.

We next give a brief overview of the control and image processing software, called ZeffControl, written within the frame of the present work. ZeffControl was designed in the Borland C++ Builder integrated development environment, employing the C++ programming

²Peripheral Component Interconnect

Parameter	Value	Default, if applicable
Number of pixels (hor. x vert.)	768×512	-
Binning	1×1 to 8×8	8×8
ADC	12 bit	-
Read-out speed	up to 12.5 Mpixel/s	4 Mpixels/s
CCD temperature	max. 40°C below ambient	0°C
Exposure time	0.005 to 9999 s	0.050 s

Table 4.1: The main operational parameters of the CCD camera used for the visible bremsstrahlung diagnostic on TEXTOR.

language. It is a multithreaded windows application (see the screenshot in Figure 4.7), and all relevant settings of the camera, different aspects of the image processing stage and of the data transfer to the central TEXTOR storage facility, can be adjusted from within the program. The concept of object-oriented programming, intrinsic to the C++ language, proved to be especially useful in designing (software) objects like a CCD camera or a data transfer object, characterized by their *properties* and *methods*, being able to inherit from a parent object.

Camera control and imaging sequence ZeffControl interfaces with the camera by means of a Windows Dynamic Link Library (DLL), provided by the company DTA, which allows controlling the camera via a set of variables and procedures. At the beginning of the day, ZeffControl initializes the camera and cools the CCD to an operating temperature of 0°C . A periodic cleaning of the CCD is carried out, to avoid the build-up of excess charge from thermal noise in the potential wells. Apart from this, the program remains idle until it receives a trigger from the TEXTOR master clock, indicating the start of a tokamak discharge sequence. Typically, the TEXTOR T-21s trigger signal is chosen, in order to allow for some time for the background activities of the camera (temperature control and CCD cleaning) to be shut down, as they otherwise would slow down the acquisition process.

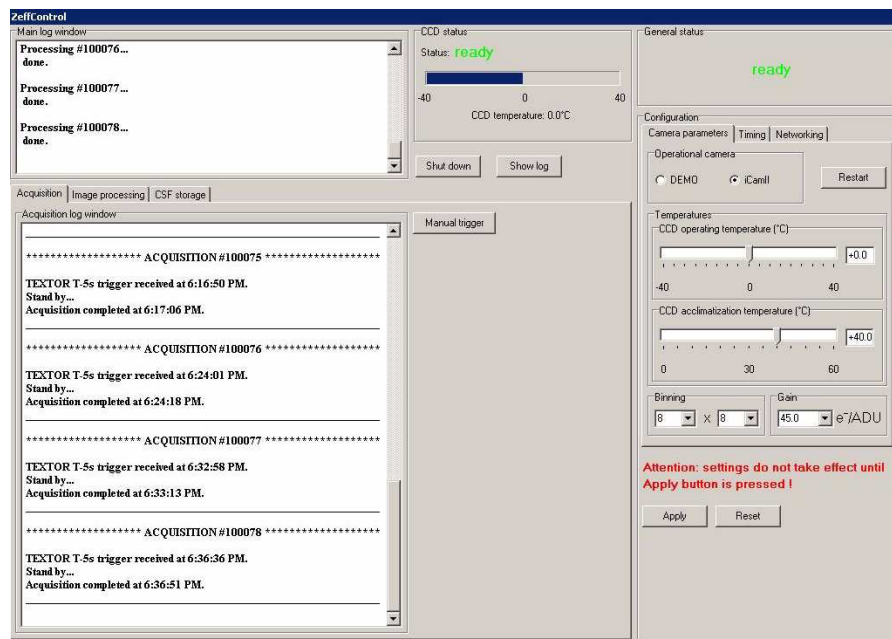


Figure 4.7: A screenshot of the interface to the control and image processing software ZeffControl of the visible bremsstrahlung Z_{eff} diagnostic on TEXTOR.

At the same time, the trigger is also sent to a programmable digital timing circuit that controls the timing of the imaging sequence. If T0 indicates the start of the discharge, which is also the start of the acquisition of the first image, then the trigger unit sends a 5 V TTL³ pulse to the workstation at T0 – 20 ms. As a result, ZeffControl instructs the camera to wait for a trigger signal. At T0 the timing unit sends this trigger signal to the camera, which consequently takes an image. As soon as the image is acquired and the image data are transferred to the workstation memory, ZeffControl returns a TTL pulse to the trigger unit to signal the successful acquisition of the image. This serves as a notice to the trigger unit that both the camera and ZeffControl are ready to take the next image. This procedure is repeated for every image during a period of ten seconds, which is sufficient to cover the duration of every TEXTOR discharge. The number of images taken in this time frame depends on the exposure time for one image, the read-out time of the CCD and the speed of the PC. Usually 200 images are taken, yielding a sampling frequency of 20 Hz. This is much higher than the sampling rate of the mirror diagnostic, which was less than 7 Hz. This 20 Hz is of course still less than what could be achieved using the previous diagnostic with seven fixed chords, but the advantage with the current diagnostic is that the number of chords is much higher. Great care was taken to ensure that the timing circuit does not suffer from small cumulative timing errors, possibly leading to a considerable de-synchronization with the TEXTOR clock. This would be a serious problem, since for the calculation of a Z_{eff} time trace, the signals for ϵ_{ff} , n_e and T_e have to be well synchronized. Otherwise, sudden changes of the plasma parameters might lead to strong artefacts in the Z_{eff} signal.

Image processing A typical image from the acquisition during a TEXTOR discharge is shown in Figure 4.8(a). One notices the light spots that are the images of the exits of the light guides. In the present configuration, 24 channels are visible. During calibration (see below), the exact position of these images on the CCD array was determined. In the image processing stage of the imaging sequence, which is performed after all images have been acquired, for every frame and for every fibre, the value (in ADU, or Analog-to-Digital Units) of a central pixel (which is in fact a bin of 64 physical pixels) in the fibre image, is compared to

³Transistor-Transistor Logic

the value of the same pixel during calibration. Care has to be taken not to choose a too long exposure time, in order to avoid saturation of the CCD. This occurs at 4095 ADU, since the dynamic range of the ADC is 12 bit. If necessary, the gain of the camera (i.e. number of photo-electrons per ADU) can be adjusted. In addition, longer exposure times naturally limit the maximum obtainable frame rate. Ultimately, for every time instant when an image was taken, a set of 24 line-integrated emissivities (in $\text{W}/(\text{cm}^2 \text{ sr nm})$ after calibration) is obtained, or, equivalently, we obtain for every channel a time trace of the emissivity. For archiving and control purposes, the raw images as well as the time traces are locally saved on disk for a limited period of time. In addition, the signals are sent to the central storage facility.

Transfer to the CSF and availability via the TWU The central database of TEXTOR is called the *Common Storage Facility*, or CSF [62], which is basically a file server for the storage of raw data. The bremsstrahlung data is transferred to the CSF using the File Transfer Protocol (FTP). An Apache server runs on the CSF and the data is meant to be accessed via the Hypertext Transfer Protocol (HTTP) using the so-called *TEC Web Umbrella* (TWU) scheme. Not only does this system provide a structured way to access all diagnostic—and processed—data from the TEXTOR experiment, but it also dispenses with the many issues involved in accessing data produced by the host of different computing platforms acquiring data at TEXTOR.

The idea of the TWU is that the data is presented to the user as a set of directories and files on the CSF web server. Any program that uses the HTTP protocol (like a web browser) can be used to explore the signals. The signals, and their properties, are addressed by URLs⁴, as is common practice for accessing web pages on the internet. However, rather than being static and readily available on the web server, these web pages are created dynamically when a request is made. Indeed, whenever signal data is being addressed, a dedicated software routine is run on the server, which creates the web page in HTML (Hypertext Markup Language)⁵. This server program was also developed in the frame of this work.

⁴Uniform Resource Locator

⁵This scheme is called the *Common Gateway Interface* (CGI).

Elimination of CCD smearing Since the shutter remains open during the entire acquisition process, light also strikes the CCD during read-out. In the specific sensor used here, the read-out involves the gradual shifting of the accumulated charge up the vertical register. This leads to each illuminated pixel being ‘smeared out’ towards the bottom of the image. This process is commonly referred to as *smearing*, and the effect is negligible at a sufficiently high ratio of exposure time to read-out time. In the present case however, in order to maximize the frame rate, a relatively short exposure time was chosen. As a consequence, the smearing effect has to be taken into account. Indeed, a typical raw image is shown in Figure 4.8(a), which suffers clearly from a considerable amount of smearing. This is even more clear in Figure 4.8(b), which is the same figure on a logarithmic scale. Vertical bands of light can be seen on the image, at positions defined by the images of the light guide exits. In this particular image, there are bands of light below *and* above each smeared channel, because the image is part of (and not the first one of) an image series, so that the previous image contributes to the smearing as well. The contribution to the value of a certain pixel due to smearing depends on the intensity of all pixels in the same column, and so smearing introduces a certain level of channel cross-talk, as well as some cross-talk between consecutive time frames. A smearing effect as seen in the images in Figure 4.8 is thus unacceptable and has to be removed.

The most obvious way to eliminate smearing is to employ a mechanical shutter in front of the camera entrance—this was the method initially used. If the shutter closes during read-out of the image, there is no more smearing effect. Unfortunately, the introduction of a shutter in the system decreases the maximum attainable sampling frequency, because opening and closing the shutter also takes a small amount of time. Even more serious, it was found that the internal electromechanical shutter that was factory-installed was malfunctioning after one year of operation, most probably due to excessive use exceeding the original design specifications. It therefore became clear that a different solution had to be found to avoid smearing.

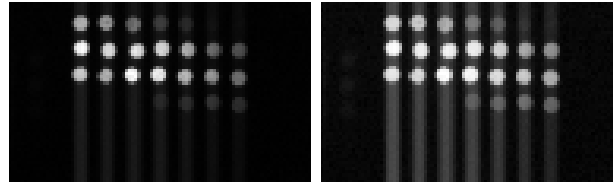
Another method for smearing removal that is sometimes used in astronomy, is by calculating the *Point Spread Function* (PSF) of the optical system, and subsequently de-convolving the image in the frequency domain. This method has the additional advantage that contributions to the PSF of the entire optical system are automatically taken into

account. Since most stars can usually be considered as point sources, this process is relatively straightforward, as one only has to observe the image of a star to calculate the PSF. For our application however, it would be required to construct a point-like source in order to assess the point spread function experimentally. We chose to avoid the additional experimental work needed to carry out this idea, and instead use a purely computational method. A first idea that might occur, is to construct an artificial point spread function that simulates the smearing effect sufficiently well. This function will depend on the ratio between the time during which smearing occurs for a certain pixel, and the exposure time of the image. In our case, there is an even simpler solution, because we know that the image is approximately dark in regions on the CCD far away from any channel image. Indeed, neglecting the difference in illumination levels between consecutive images, it can be seen that every pixel in a certain column receives the same amount of smearing S . For every column, S can be approximated by averaging the difference between signal and dark current over a range of pixels in a dark region in the same single column. Then, S is subtracted from the value of every pixel in this column. This procedure leads to satisfactory results, as can be seen in Figure 4.8(c), where an image is shown that was de-smearred using this approach. This also validates the approximations made.

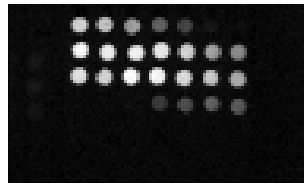
Additional photomultiplier measurements As mentioned before, seven channels can be illuminated that were from the beginning of the design intended to be equipped with photomultipliers, which can achieve a considerably higher sampling rate than the camera system (20 kHz versus 20 Hz). With this system, it becomes possible to study transient phenomena. In addition, this system serves as a back-up for the camera system. Up to this moment, four photomultipliers have been connected.

Reflections within the vacuum vessel We conclude this section on the upgraded bremsstrahlung diagnostic with a note on the issue of reflection of plasma radiation on surfaces inside the vacuum vessel.

When the first version of the upgraded diagnostic was tested, it quickly became apparent that several channels suffered from interference by radiation entering the lines of sight after reflection on metallic surfaces (flanges, liner, etc.) within the vessel. In fact, the reflections



(a) A typical CCD image, (b) The same image on a logarithmic scale.



(c) The image after de-smearing, on a logarithmic scale

Figure 4.8: Removing the smearing effect in the CCD images.

caused a diversion of the line-integrated values that was so important, that the reconstruction of a bremsstrahlung emissivity profile turned out to be not feasible.

Following this observation, a small viewing dump made out of graphite tiles was installed in december 2004 on some of the surfaces causing the most prevalent reflections. However, further tests showed that also this modification was inadequate, and that a larger viewing dump was necessary. The installation of such a viewing dump was accomplished in the summer of 2005. This viewing dump is fabricated from sandblasted stainless steel. Measurements indicate that the viewing dump is effective.

4.1.2 Calibration

The calibration of a measurement system can be a complex issue, and if performed with insufficient care, can lead to large deviations in the measurement results. When considering the calibration of any multichannel diagnostic, one has to discriminate between the *relative* (channel-to-channel) calibration of the channels with respect to one another, and the *absolute* calibration of the channels in terms of some system of

physical units.

In many studies employing some estimate for Z_{eff} , a line-averaged Z_{eff} value along a certain line of sight already serves as a good indication for the general impurity levels in the plasma. Moreover, frequently one is rather interested in the *relative* variations of Z_{eff} , be it a line-averaged or a local value, as a results of certain changing plasma conditions. For example, in studies of impurity screening under the effect of the DED on TEXTOR [63–65], one is interested in a possible change in Z_{eff} when the DED current is switched on. If indeed only relative Z_{eff} variations in the global plasma are studied, then in principle a calibration of the channel that is used for the determination of the line-averaged Z_{eff} , is not needed. If, however, the interest lies in a more localized estimate of impurity levels, then one has to perform some sort of inversion of the line-integrated bremsstrahlung emissivity measurements. In this case, a relative calibration of the channels is necessary. Finally, if some knowledge of the absolute value of the impurity concentration is needed, then the system for bremsstrahlung emissivity measurements requires an absolute calibration. This is then also true for the diagnostics that deliver the density and temperature data, but we will not discuss this here.

4.1.2.1 Integrating spheres

In the case of the visible bremsstrahlung diagnostic that we are discussing here, a relative calibration of the channels requires a source of constant emissivity, and an absolute calibration requires a source of known emissivity. The *integrating sphere* is a device that provides both at the same time. An integrating sphere is coated on the inside with a specialized diffusing and highly reflective material (typically over 95% reflective in the wavelength region of interest). A lamp irradiates the sphere surface, and the sphere has one or several relatively small exit ports to admit the radiation to leave the sphere. The lamp is not located on a line of sight towards any opening, so that the light exiting the sphere has undergone in general many reflections. This can be realized in practice by installing baffles inside the sphere, or by locating the lamp in a satellite sphere. The spectral sphere surface radiance as seen by an external observer looking inside the sphere, can then be written approximately as an infinite power series that converges to

$$L_{s\lambda} = \frac{\Phi_{i\lambda}}{\pi A_s} \frac{\rho(\lambda)}{1 - \rho(\lambda)(1 - f)} \quad (\text{W} / (\text{m}^2 \text{ sr})),$$

where:

- $\rho(\lambda)$ = spectral reflectance of the sphere surface,
- $\Phi_{i\lambda}$ = input spectral radiant flux,
- A_s = total sphere surface,
- λ = wavelength.

The radiance losses through the exit ports are taken into account via the port fraction f . A tungsten halogen lamp is often used, providing a continuous spectrum free of line emission. In that case, the spectral input flux $\Phi_{i\lambda}$ can be expressed in the familiar blackbody form. The integrating sphere is approximately a Lambertian radiator, radiating isotropically. The luminance across the exit port of a non-ideal sphere varies slightly. The deviations are however too small to be considered in a diagnostic system where sources of far greater errors inevitably exist.

The integrating sphere used for the calibration of the bremsstrahlung system was a Labsphere US-060-SF model with a diameter of six inch and coated with Spectrafect. It is shown during calibration inside the TEXTOR vessel in Figure 4.9.

4.1.2.2 Calibration procedure

The calibration was performed with the same camera settings as during standard operation. To completely eliminate smearing however, the camera exposure time was set to a few seconds, while the external shutter was opened by means of an external TTL signal during a few milliseconds: the effective exposure time for each calibration image.

The geometry of the chords was first determined by illuminating the fibre head at the camera side by a simple lamp, and the location of the light spots from the fibres was measured in the focal plane of the objective at the TEXTOR side, which is the equatorial plane of the torus. The field of view is shown in Figure 4.4. The integrating sphere was then mounted on a small optical bench inside the tokamak vessel, so that, starting from the known chord geometry, the sphere could be moved such that the line of sight of every fibre, one after another, was pointed towards to exit port of the sphere (Figure 4.9). If the by the camera measured emissivities from the sphere, in ADU, are labeled by

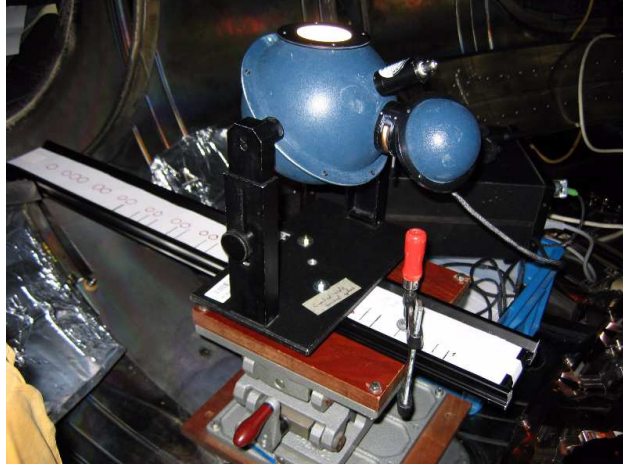


Figure 4.9: The integrating sphere mounted inside the TEXTOR vessel. The viewing geometry of the lines of sight has been indicated on a strip of paper.

$L_{s,1}, \dots, L_{s,24}$, and the measured line-integrated emissivities from the plasma, also in ADU, are designated by $L_{p,1}, \dots, L_{p,24}$, then

$$L_{ff,i} = \frac{E_s}{L_{s,i}} \frac{T_c}{T_a} L_{p,i}, \quad i = 1, \dots, 24, \quad (4.1)$$

where

- $L_{ff,i}$ = physical line-integrated bremsstrahlung emissivity along chord i , in $\text{W} / (\text{cm}^2 \text{ sr nm})$
- E_s = spectral emissivity of the lamp at λ_{ff} ,
- T_a = exposure time during acquisition,
- T_c = exposure time during calibration.

In the following, we will define:

$$c_i \equiv \frac{L_{s,i}}{E_s} \frac{T_a}{T_c},$$

and call the c_i the *calibration factors*, so that

$$L_{ff,i} = \frac{L_{p,i}}{c_i}, \quad \forall i.$$

4.1.2.3 Calibration difficulties

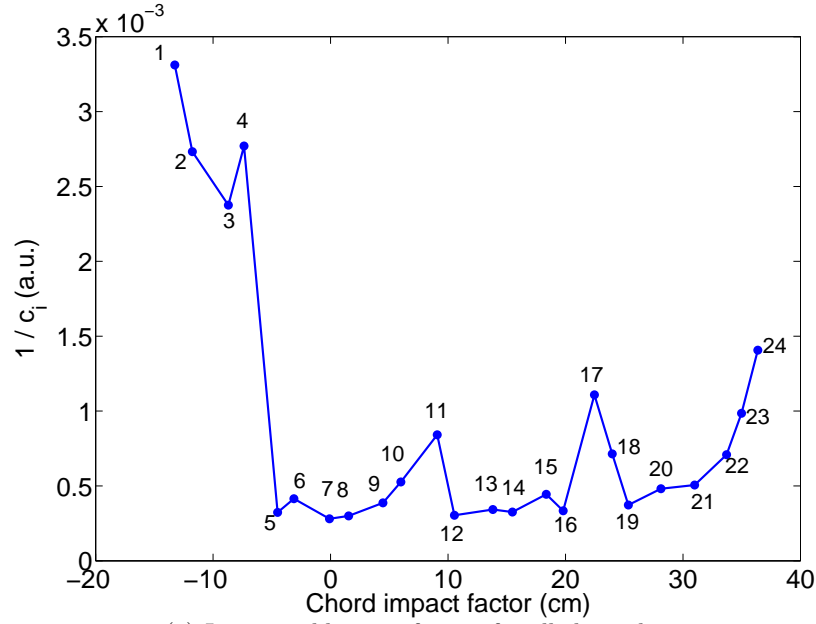
The calibration of a spectroscopic diagnostic using a reference source of emissivity, as outlined above, is very sensitive to a host of possible sources of error, which are very difficult to identify. As soon as the calibration factors for the Z_{eff} diagnostic were applied to the first sets of measurements, it became clear that their values, relative to each other, were not realistic. This fact becomes apparent when considering the graphs in Figure 4.10. In Figure 4.10(a) the inverse calibration factors are shown for the 24 channels. The variation in calibration factors is mainly due to differences in fibre quality and lens throughput for different channels. Figure 4.10(b) displays a typical line-integrated emissivity profile resulting from the application of this calibration. There are a few similarities between both curves, the most striking being visible in channels 11 and 17. This observation implies that the measurement conditions during the measurement of the emissivity from the integrating sphere, are not equivalent to the conditions during the measurement of plasma emissivity, and that this behavior is channel-dependent. The reason for this calibration problem has up to now not been found. Successively, CCD smearing and in-vessel reflections have been presumed to be responsible, but these effects have all been neutralized now, while the problem persists. A further cause might be that the integrating sphere does not radiate perfectly isotropically, but this possibility has not been explored as yet.

One last remark: in Figure 4.10 the channels are numbered from 1 to 24 starting from the TEXTOR High Field Side. This is a convention that we will follow throughout the rest of this work.

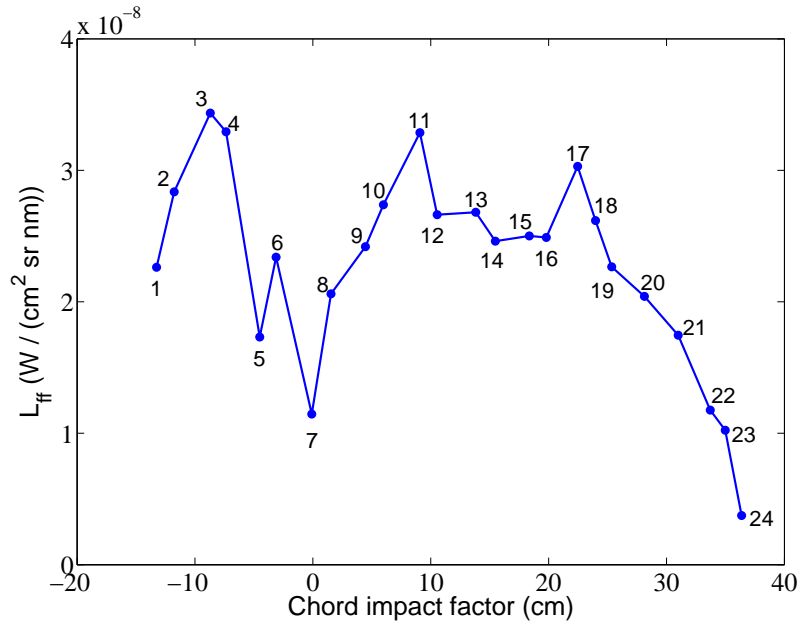
4.1.2.4 Relative calibration estimation from profile consistency

In order to estimate the correct relative calibration, an alternative procedure was devised in the context of the current work, as described below. The method involves using the plasma itself as a calibration light source [66].

Suppose two channels, say A and B , measure along the same physical line of sight through the plasma. This results in two measurements with the camera, m_A and m_B , respectively, both expressed in ADU. Either measurement should, apart from any measurement error, lead to the same line-integrated emissivity L_{ff} :



(a) Inverse calibration factors for all channels.



(b) A typical line-integrated emissivity profile following from the above calibration.

Figure 4.10: Influence of an erroneous calibration on the line-integrated emissivity profile. The data points are interconnected to facilitate recognizing the similarities in both curves.

$$L_{\text{ff}} = \frac{m_A}{a} = \frac{m_B}{b}. \quad (4.2)$$

Here, a and b are the calibration factors associated with channels A and B , respectively. If a is fixed to an arbitrary value, then the relative calibration of this two-channel system can be calculated. This procedure can be repeated for the other channels, and eventually the complete relative calibration of the 24-channel system can be assessed.

In practice, the question rises how to feed the same line-integral L_{ff} to different channels. At this point, we have to assume that the line-integrated emissivity profile is sufficiently smooth, in the sense that there are no ‘sharp peaks’ in the profile. To be more exact, if we interpolate the physical line-integrated emissivities as a function of chord intercept b , with a smooth curve $f(b)$, then d^2f/db^2 should be sufficiently small in absolute value. This requirement is fulfilled in general. In such cases, the calibration factors can be estimated in a way that differs slightly from the technique described above, but which follows the same principle. The line-integrated bremsstrahlung emissivities along the 24 channels are measured in two discharges with very similar plasma parameters (which we assume to be identical), but with a relative shift H in horizontal plasma position. In practice a few cm is sufficient, while H is taken positive when the shift is towards the HFS of the machine. In addition, if the plasma shift is not taken too large, the plasma-wall interaction will not change to the extent that there will be an appreciable change in the bremsstrahlung emissivity throughout the plasma. This results in two sets of measurements (m_i , respectively n_i) along two chord fans with different intercepts on the equatorial plane, relative to the plasma⁶, (b_i , resp. $b_i + H$, $i = 1 \dots 24$), but with the same calibration factors c_i . Every measurement m_i , resp. n_i , corresponds with a calibrated line-integral $M_i = \frac{m_i}{c_i}$, resp. $N_i = \frac{n_i}{c_i}$. The M_i are now interpolated as a function of intercept using a relatively smooth curve $f(b)$. Alternatively, we may use a low order polynomial fit, which is less biased towards individual points $M_i(b_i)$. The latter is particularly useful in the case where the number of channels within the field of view is relatively low. In both cases, to a good approximation, the following equalities should hold:

$$N_i = f(b_i + H), \quad i = 1, \dots, 24. \quad (4.3)$$

⁶Naturally, the intercept relative to the machine does not change under a plasma shift.

These equations simply express the fact that in both cases the same line-integrated profile was measured. To avoid extrapolation of the emissivity profile, and depending on H , the equations for some outer channels are not taken into account, and the corresponding calibration factors are not calculated. In the case of linear interpolation of the line-integrated emissivity profile, the calculation of the relative calibration factors amounts to the solution of a system of linear equations. The shift H should then not be too large, otherwise the system may become unsolvable. Equation (4.3) is the equivalent of (4.2), and again the only calibration factors that satisfy all equations, are the real ones. In fact, again, with the here described method, the calibration factors can only be found up to a certain factor. Indeed, if a set of calibration factors c_i is found that satisfies (4.3), then it is easy to see that also the set $a \cdot c_i$ will satisfy these equations, for an arbitrary factor a . Although the relative calibration on its own already allows the reconstruction of relative Z_{eff} profiles on an arbitrary scale, an absolute calibration is required for the assessment of absolute impurity concentration levels. The full absolute calibration can be fixed by the calibration of a single channel with an integrating sphere. This requires only a reduced experimental set-up, as compared to the calibration of all channels with the sphere.

Another way to see that the relative calibration is determined uniquely by (4.3), is the following. From an artificial (calibrated) and relatively smooth line-integrated profile $L_{\text{ff},1,i}$, $i = 1, \dots, 24$, the corresponding ‘measurements’ (in ADU) on the CCD camera are calculated, assuming an arbitrary set of calibration factors c_i , resulting in $m_i = c_i \times L_{\text{ff},1,i}$. Then, a plasma shift of 1 cm towards the high-field-side is simulated by interpolating the $L_{\text{ff},1,i}$ at the shifted intercepts $b_i + H$, yielding the line-integrals $L_{\text{ff},2,i}$ and a new collection of measurements $n_i = c_i \times L_{\text{ff},2,i}$. Now, one of the calibration factors (here c_{10}) is decreased by 20 percent, and the new set of calibration factors is called d_i . With this adapted calibration, a set of ‘calibrated’ line-integrals is calculated: $M_i(b_i) = \frac{m_i}{d_i}$ and $N_i(b_i + H) = \frac{n_i}{d_i}$. Figure 4.11 shows a plot of both the M_i and the N_i , as a function of impact factor. Naturally, both line-integrated profiles are identical (barring the shift), except at their respective channel 10, generating the obvious peak in the profile. However, the important point is now that the profiles are not entirely overlapping, because channel 10 is in both cases associated with a different intercept (b_{10} versus $b_{10} + H$). Hence, the statement

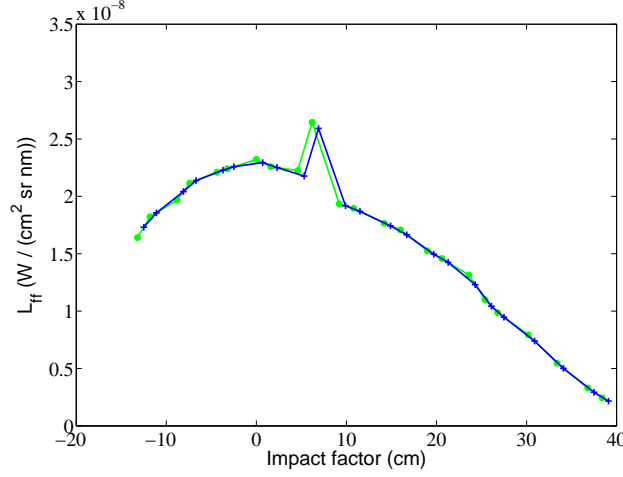


Figure 4.11: Simulated calibrated line-integrated emissivity profiles, original (blue line) and after the plasma shift has occurred (green line). The calibration factor of channel 10 has artificially been altered from the true one.

follows that the calibration that minimizes the difference between the two profiles, is the correct one (up to a factor).

Figure 4.12 gives a visual representation of equations (4.3). The geometry is shown of a line of sight, viewed by channel k , under a plasma shift H towards the high field side. In the figure, the plasma is depicted stationary, and instead an equivalent shift of the line of sight is shown. The blue line represents the line of sight C_k (intercept b_k with respect to the centre of the plasma column) viewed by channel k before the shift has occurred. The black line is a second, adjacent line of sight C_l viewed by another channel l . These channels view a physical line-integrated emissivity M_k , resp. M_l . The green line depicts the line of sight viewed by channel k , after the shift has occurred (intercept $b_k + H$, relative to the plasma). Channel k now views an emissivity N_k . Next to a shift in intercept, there is also a shift of observation point (A to B), relative to the plasma. The red line represents the line of sight for a fictitious channel, with associated emissivity $f(b_k + H)$, resulting from the fitting or interpolation process of the initial set of line-integrated emissivities M_i . The intercept for the red chord is the same ($b_k + H$) as for the green chord. Now, according to (4.3), N_k

and $f(b_k + H)$ should be equal. However, it can be seen in Figure 4.12 that this is in reality not entirely true. On the other hand, it is also clear that in the present geometry, the difference δ in emissivity viewed by the red and the green chord, is relatively small. Simulations with a standard radial bremsstrahlung profile show that δ is typically a few percent or less of the emissivity along the shifted (green) chord. Therefore, it is allowed to neglect the shift of observation point relative to the plasma. Nevertheless, a slightly better approach may still be to rotate the fibre optics viewing inside the plasma, instead of shifting the plasma itself. In summer 2006, some adjustments were made to the mechanical support of the fibre optics at the TEXTOR side, in order to allow a more precise and controlled rotation. Tests with this calibration approach will be performed in fall 2006.

The artificial data set that was used above was again employed to test the relative calibration method, except that now a plasma shift of 2 cm towards the HFS was imposed. The calibration factors c_i were assumed to be the real physical ones. An arbitrary set of calibration factors $d_i \equiv 100$, different from the c_i , were used as an initial state. Figure 4.13 shows the result of the method, indicating that the assumed calibration was approximated very well.

Finally, the calibration method was applied using real data from two TEXTOR discharges (#99431 and #99430, both at 3 s) with very similar plasma conditions, and with a horizontal shift of 1 cm towards the HFS (from #99431 to #99430). The results are presented in Figure 4.14. The line-integrated bremsstrahlung emissivity profile for both discharges is plotted against chord intercept b , relative to the plasma. The profiles were calibrated by the calibration factors that were calculated via linear interpolation of the starting profile. There is a good correspondence between the two profiles, implying that the equations (4.3) are well satisfied. The outermost channels at both sides were discarded, since the corresponding calculated calibration factors were not satisfactory. The absolute calibration was assessed from the calibration of a single channel using the integrating sphere, viewing along the normal onto the sphere exit plane. The thus determined calibration yields physically acceptable radial Z_{eff} profiles over most of the plasma cross-section (see Section 4.1.3). Future simulations and experiments are planned to estimate the error bars of a set of calculated calibration factors.

The here outlined relative calibration method has been derived be-

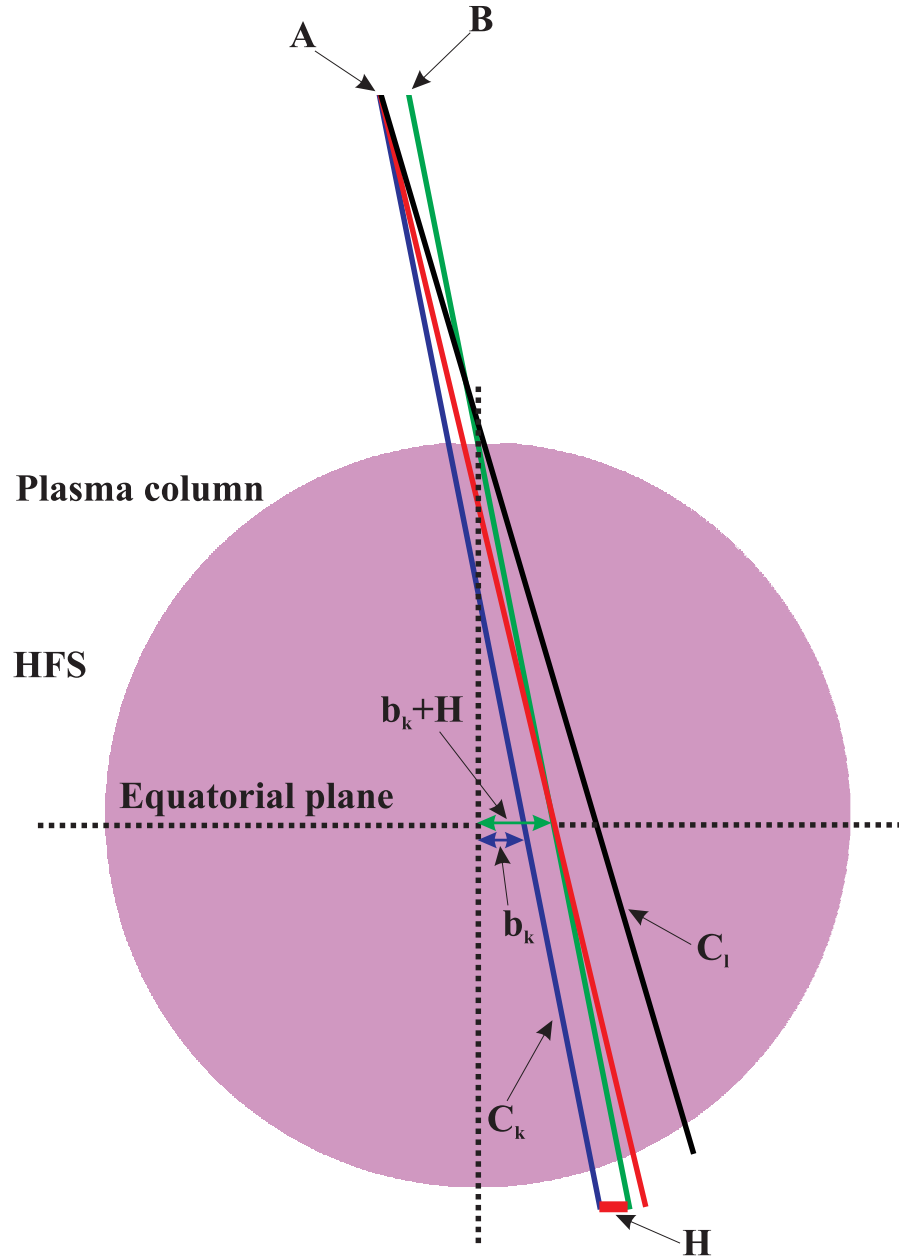


Figure 4.12: The geometry of a chord during a horizontal plasma shift.

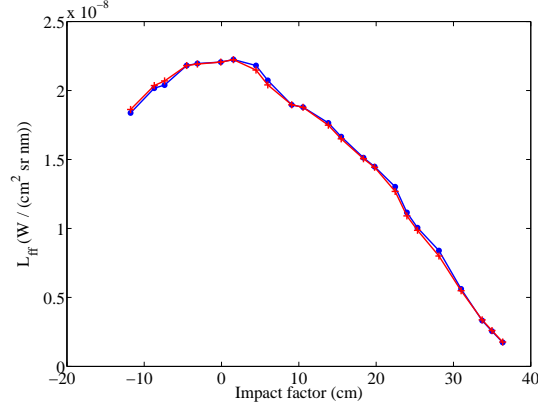


Figure 4.13: A simulation of the relative calibration method based on (4.3). The blue curve represents the original artificial line-integrated bremsstrahlung emissivity profile, the red curve is the profile based on the calculated calibration. The profile based on the calculated calibration was rescaled and shifted back to match the original profile.

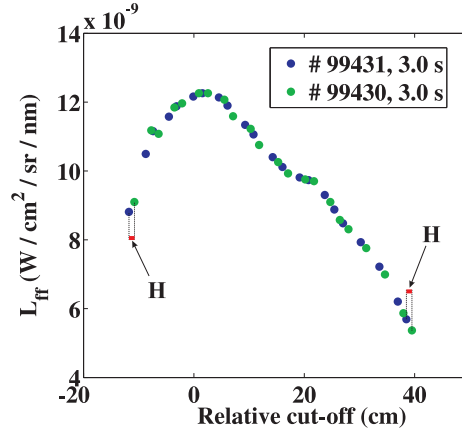


Figure 4.14: The calibrated line-integrated bremsstrahlung profiles resulting from the requirement of profile consistency under a horizontal plasma shift H . H is indicated in the figure for two channels.

cause of the existing problems with the calibration using the integrating sphere. However, the technique is of general applicability, and provides a simple and self-consistent way for the relative calibration of any multichannel spectroscopic diagnostic⁷. It has several advantages over a calibration using a dedicated light source. Indeed, this method of relative calibration can in principle be carried out at any time, without requiring access to the tokamak vessel itself or the establishment of a calibration set-up. The fibre optics also do not have to be removed from the machine, avoiding changes in alignment of the fibres. Since the plasma acts as a calibration light source, no specialized light sources are needed for the relative calibration. This is a considerable advantage, especially in the case of spectroscopy outside the visible range. In addition, this method provides a valuable way for the check or optimization of an existing relative calibration. Finally, in situ calibration procedures will be very difficult in future fusion devices, at least on a frequent basis. Therefore, calibration methods as the one described here, using the plasma itself as a reference, will only gain in importance.

On TEXTOR, an upgraded diagnostic for carbon spectroscopy has recently been installed. The brilliance of a CIII and a CV line is measured in the ultraviolet, each along nine chords. The relative calibration of the channels requires a UV light source emitting approximately at the wavelengths observed for CIII (229.6 nm) and CV (227.1 nm). The brilliance of the CV line throughout the plasma does not change appreciably if the plasma is shifted horizontally over a small distance. Therefore, the method for relative calibration of the channels based on requirements of profile consistency, can also be carried out for the relative calibration of the channels measuring the CV line. In contrast, the CIII line brilliance does change when shifting the plasma, due to a change in plasma-wall interaction. Nevertheless, we can perform a relative calibration of the CIII channels as well, if we tune the associated spectrometer to the CV line. This calibration is planned for autumn 2006.

4.1.3 Bremsstrahlung emissivity profile reconstruction

From the calibrated line-integrated bremsstrahlung emissivity measurements, recorded by the camera, radial bremsstrahlung profiles can be

⁷One requirement, however, is that shifting the plasma has little influence on the emissivity throughout the plasma of the observed radiation.

reconstructed⁸. Traditionally, on TEXTOR this is done using an Abel inversion [51]. This involves the comparison of the experimental line-integrals with calculated integrals for an emissivity profile $\epsilon_{\text{ff}}(r)$ that is expressed by a coefficient vector \mathbf{x} in terms of a set of basis functions. \mathbf{x} is required to satisfy $\mathbf{y} = M\mathbf{x}$, with \mathbf{y} the set of experimental line-integrals, and M the so-called *geometry matrix*⁹. Here, \mathbf{x} is estimated by minimizing the squared differences (residuals) between the calculated and experimental line-integrals.

It is commonly known that the reconstruction of profiles from a limited set of line-integrals, is a difficult operation. One of the main reasons is that the inversion problem is severely ill-conditioned. Condition numbers¹⁰ typically vary between 10^2 and 10^6 . Every uncertainty on the line-integrals is usually magnified in, and can propagate through the profile. Some sources of uncertainty on the line-integrals were already highlighted in Section 3.3. For example, for the inversion poloidal symmetry has to be assumed since there is only a single observation point. A serious issue is posed by the unknown contributions to the continuum radiation near the plasma boundary, leading to an overestimation of the edge Z_{eff} value, which is particularly emphasized in the case of Abel inversion. In addition, such an edge contribution can propagate towards the centre of the profile [67].

Therefore, in the frame of the current work, the inverse problem of profile reconstruction was approached using Tikhonov and Maximum Entropy regularization [59]. The Tikhonov regularized solution \mathbf{x}_α to the inverse problem $\mathbf{y} = M\mathbf{x}$ is the minimizer of the functional (see e.g. Ref. [68])

$$F_\alpha(\mathbf{x}) = \|M\mathbf{x} - \mathbf{y}\|^2 + \alpha\|\mathbf{x}\|^2. \quad (4.4)$$

So, in addition to the usual inversion problem, Tikhonov regularization involves controlling also the norm of the solution. This prevents wild oscillation of the solution \mathbf{x} due to overfitting and a near to singular geometry matrix. The regularization parameter α was determined using the L-curve method [69], which allows to minimize the data misfit, while preventing overfitting.

Similarly, the Maximum Entropy regularization involves the maximiza-

⁸The four channels coupled to a photomultiplier do not provide sufficient information for reliable profile reconstruction.

⁹See also Section 6.5 for some of the details behind profile reconstruction.

¹⁰The *condition number* of a matrix is defined as the ratio between the largest and smallest non-singular element.

tion of the functional

$$F_\beta(\mathbf{x}) = -\|M\mathbf{x} - \mathbf{y}\|^2 + \beta S(\mathbf{x}), \quad (4.5)$$

with

$$S(\mathbf{x}) \equiv \sum_i x_i - m_i - x_i \log \frac{x_i}{m_i} \quad (4.6)$$

the entropy of \mathbf{x} (with components x_i), and m_i are the components of a default model vector [70]. The latter allows the inclusion of any available prior knowledge about the profile into the analysis. Until now, all m_i were taken equal. Put simply, the rationale for maximizing the entropy of the solution is that this minimizes the assumptions one makes about \mathbf{x} that are not supported by the data.

Figure 4.15 demonstrates the reconstruction of a Z_{eff} profile in a TEXTOR discharge using Abel inversion, and Tikhonov and Maximum Entropy regularized inversion. The calibration used is the one obtained from profile consistency requirements in Section 4.1.2.4. For the regularized inversion, a set of B-splines was used as a basis. In the three cases, the Z_{eff} value is unrealistically high near the plasma boundary, which is most likely due to the various edge continuum contributions apart from bremsstrahlung. This inconsistency is minimized using the Tikhonov and Maximum Entropy regularization. The sensitivity of the central Z_{eff} to uncertainty in the edge continuum, in the case of Tikhonov and Maximum Entropy regularization, is currently under study.

4.1.4 Z_{eff} calculation

The bremsstrahlung emissivity measurements made by the diagnostic described here, are used for the calculation of Z_{eff} . This requires electron density and electron temperature data, which are obtained from interferometry and from Electron Cyclotron Emission, respectively (see e.g. Ref. [29]). A line-averaged Z_{eff} is routinely calculated and Z_{eff} profiles can be reconstructed on demand.

4.1.4.1 Line-averaged Z_{eff}

A line-averaged Z_{eff} value is calculated from the line-integrated bremsstrahlung emissivity, measured by the camera along one centrally pointed chord. If the plasma parameters are assumed to be constant on a

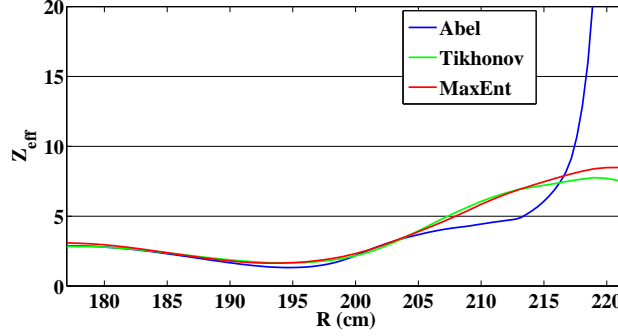


Figure 4.15: Reconstructed Z_{eff} profiles using Abel inversion, Tikhonov and Maximum Entropy regularized inversion for TEXTOR discharge #99434 at 1.9 s. The Z_{eff} overestimation at the plasma edge is the largest for the Abel inverted profile.

flux surface (poloidal symmetry), this line-averaged Z_{eff} can be defined as [71]

$$\bar{Z}_{\text{eff}} \sim \frac{\int_{-a}^a \epsilon_{\text{ff}}(r) ds}{\int_{-a}^a \bar{g}_{\text{ff}}[T_e(r)] n_e^2(r) T_e^{-1/2}(r) ds},$$

where a is the plasma radius, and s denotes a length parameter along the line of sight. The calculation of \bar{Z}_{eff} is carried out automatically on an intershot basis, for every time point at which bremsstrahlung data is recorded by the camera system. A code was written in the context of the current work to perform the calculations. The resulting signal for the line-averaged Z_{eff} is available on-line through the so-called TPD, which is to be consulted via the TWU system as well. A similar calculation is performed for one centrally viewing chord coupled to a PM. Before the operation of this diagnostic, no Z_{eff} value was available on-line.

4.1.4.2 Z_{eff} profile calculation

From the bremsstrahlung emissivity profiles, and using also profiles for n_e and T_e , Z_{eff} profiles can be calculated. The profiles calculated within the frame of the current work have aided in several studies, about the TEXTOR Radiative Improved mode [72, 73] and about impurity transport under the influence of the DED on TEXTOR [63–65, 74].

4.1.4.3 Propagation of edge continuum contributions in Z_{eff} profiles

As mentioned before, (unknown) contributions to the edge continuum can propagate towards the centre of the reconstructed Z_{eff} profile. A possible consequence of this effect is now illustrated on the basis of a TEXTOR discharge where carbon was the dominant impurity. The Dynamic Ergodic Divertor was applied during 1.5 s, in the static $m/n = 6/2$ mode, leading to the ergodization of the edge plasma. In Figure 4.16 the time traces of the CIII and CV line intensity are shown. These intensity signals were obtained from a diagnostic for carbon spectroscopy that periodically scans a set of nine chords throughout a poloidal cross-section. A clear increase in CIII intensity can be noticed, with a simultaneous decrease in CV intensity during DED. This decrease can also be seen on the CVI intensity signal from CXS. An explanation of these carbon signals in terms of a plasma decontamination, or a screening effect, induced by DED, has been given in Ref. [63]. Thus, one would expect a decrease of Z_{eff} in the plasma centre. However, according to Figure 4.16, the Z_{eff} value in the central plasma appears to be unaffected by the action of DED. On the other hand, there is a clear increase of the reconstructed Z_{eff} at the plasma edge, which is reflected in the increase of the line-averaged Z_{eff} . This can also be seen in Figure 4.17 in the full Z_{eff} profile (obtained using Tikhonov regularized inversion) at two different times before and during DED. The constancy of the central Z_{eff} under the action of DED can be related to a large increase of the edge continuum radiation when DED is switched on. Indeed, such an increase can propagate towards the centre of the reconstructed profile, where it may conceal a relatively small *decrease* in Z_{eff} . This would be consistent with the other spectroscopic data. Thus, at present, relatively small changes in the central Z_{eff} can not be resolved when there is a simultaneous large edge variation. This represents a limit of the present diagnostic for bremsstrahlung emissivity measurement on TEXTOR. It is possible that this situation can be improved using a wider viewing angle, and this will be the subject of further investigation.

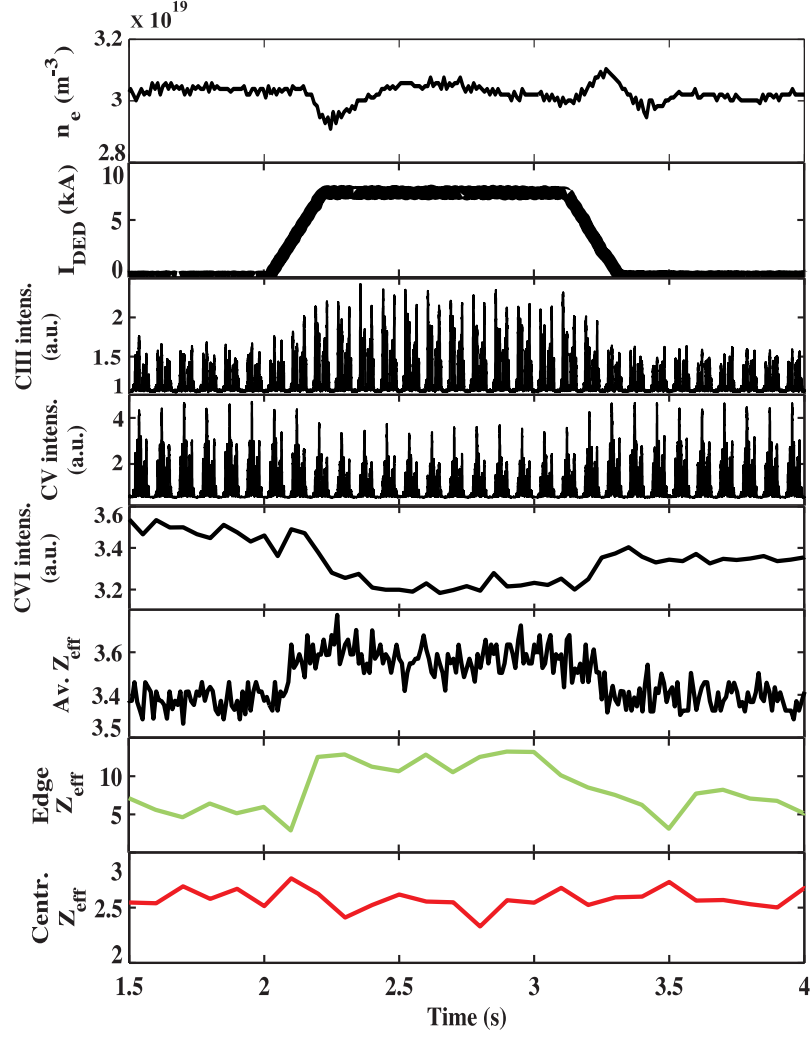


Figure 4.16: Time traces of respectively: electron density, DED current, CIII, CV and CVI line intensity, line-averaged Z_{eff} and edge and central Z_{eff} for TEXTOR discharge #99433. From the carbon line intensity signals, a decontamination or screening effect due to DED can be noticed. However, the central Z_{eff} appears to be unaffected by DED.

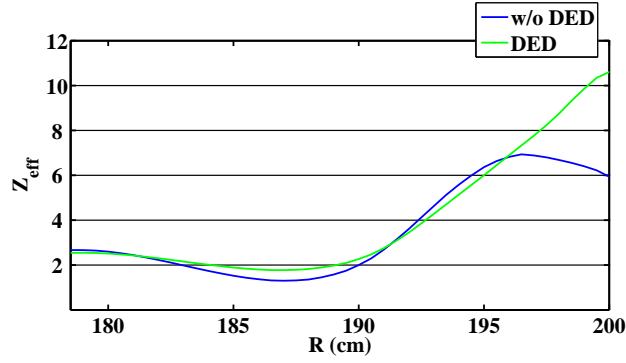


Figure 4.17: Influence of DED on the Z_{eff} profile for TEXTOR discharge #99433.

Quantity	Relative error
n_e	5%
T_e	10%
ϵ_{ff}	20%
n_C	35%
Continuum Z_{eff}	25%
CX Z_{eff}	20%

Table 4.2: The estimated relative errors on several local plasma quantities in JET or TEXTOR plasmas. The errors are generally both of a statistical and systematic nature.

4.2 The CXS diagnostic on JET

Since we will mainly work with CXS signals from JET we will give a very short overview of the JET CX diagnostic. Further information can be found e.g. in Refs. [3] and [53]. There is a vertical system consisting of one sight line, and two multichord horizontal systems. Several spectrometers, equipped with linear (for the vertical system) or two-dimensional (for the horizontal systems) CCD cameras, are coupled to the plasma via fibre optics. By using a periscope to view inside the plasma, the lines of sight can be made more or less tangential to the magnetic flux surfaces at the intersection with the neutral beam. On JET, the CX lines of sight are aligned on two PINIs¹¹ of the neutral beam in octant 8. The typical time resolution for the multichord systems is 50 ms. The spatial resolution is roughly 7 cm, depending amongst others on the number of PINIs used. Apart from the possibility to monitor carbon or beryllium and deuterium, other spectrometers have been introduced in order to measure CX lines from puffed neon and argon [75, 76]. In this work, mostly fully stripped carbon will be considered, which is usually monitored by CX at the strongest CVI transition in the visible, namely $n = 8 \rightarrow 7$ at 5290.5 Å. Since JET is not equipped with a dedicated diagnostic for space-resolved visible bremsstrahlung measurements, we will use the background level of the CX spectra in the neighbourhood of 529 nm as a bremsstrahlung emissivity measurement for use in the experiments in Chapter 6. All CX data used in that analysis were obtained from the JET KS5A instrument, with its associated observation port in octant 7.

4.3 Error estimates

For all measured quantities that are of importance in this work, error estimates are proposed by the respective diagnosticians. However, and this is the case for many measured plasma quantities in general, often the error estimates have been established by empirical arguments, requiring a lot of assumptions, and by rough approximation. The rigorous estimation of error levels for the quantities of interest here would require a dedicated approach. Therefore, we have relied on the error estimates that are in common use among diagnosticians [51, 53]. However, it is not always clear whether cited error bars are meant to

¹¹Positive Ion Neutral Injector

represent statistical uncertainty, or are rather of a combined statistical and systematic nature. Furthermore, if the errors should be interpreted as statistical errors, do they signify a single standard deviation, three standard deviations, or still something different? In most cases, the errors should be understood to be caused both by statistical and systematic effects. But also here one can pose the question with what probability the real quantity should lie within the indicated error bars.

We nevertheless would like to mention here a few commonly used relative errors on several local plasma quantities that are of interest in this work, without worrying for the moment about any of the issues posed above. The errors are indicated in Table 4.2, applying roughly both to TEXTOR and JET. The relative error for the CX Z_{eff} depends on the ratio of the carbon concentration to the CX Z_{eff} itself, but we have taken a typical value here. It should also be mentioned that the error bars can in general increase above the values stated here in the vicinity of the edge region of the plasma.

Les questions les plus importantes de la vie ne sont en effet, pour la plupart, que des problèmes de probabilité.

Pierre-Simon Laplace

He uses statistics as a drunken man uses lamp posts—for support rather than illumination.

Andrew Lang

The law that entropy always increases—the Second Law of Thermodynamics—holds, I think, the supreme position among the laws of physics. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell’s equations—then so much the worse for Maxwell’s equations. If it is found to be contradicted by observation—well, these experimentalists do bungle things from time to time. But if your theory is found to be against the Second Law of Thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.

Sir Arthur Eddington

The shortest path between two truths in the real domain passes through the complex domain.

Jacques Hadamard

5

Bayesian probability theory and Bayesian computation

The goal of this chapter is to give an overview of the Bayesian methods that we will use in the next chapter for the estimation of a Z_{eff} value that is consistent both with measurements of bremsstrahlung emissivity and CX impurity density. We will not go in too much detail, although we will occasionally spend some more time on those concepts that particularly help understanding our motivation for using the various Bayesian techniques. We start with a brief account on Bayesian probability theory.

5.1 Bayesian probability theory

5.1.1 Bayesians versus frequentists

Although commonly used in everyday speech, the word ‘probability’ is subject to several possible interpretations, which are even up to today much debated on. About three hundred years ago, people started to think seriously about the question how to reason in situations where there is no (approximate) certainty. While the emerging field of probability theory could be applied with success to calculate probabilities for various outcomes in e.g. games of chance, the question remained how to assess the probability of a certain value for not directly observed quantities, given a set of related measurements, which is actually the inverse problem. This question was first posed by James Bernoulli in 1713, and it was solved by the reverend Thomas Bayes in 1763 (in a paper posthumously published by a friend). The present-day form of the theorem, which is named after him (see below), is actually due to Laplace (1812). Not only did Laplace rediscover Bayes’ theorem, in much more clarity than Bayes did, but he also applied it in solving problems in celestial mechanics and medical statistics. Despite the many successes, Laplace’s development of probability theory was rejected by mid-nineteenth century mathematicians. Indeed, according to Bernoulli, Bayes and Laplace, a probability represented a degree of belief, or plausibility: how much they thought that something was true, based on the available data. This was later deemed too vague an idea, and probability was redefined as a long-run relative frequency with which an event occurred, given (infinitely) many repeated trials. Since frequency can be measured, probability was now regarded as an objective measure of uncertainty.

The frequency definition of probability is generally used by the so-called *frequentist school*. Although this is the definition that is commonly used in introductory textbooks on probability theory, it has a number of difficulties attached to it, as compared to the definition of probability as a degree of belief, adopted by the *Bayesian school*. The most obvious difficulty is the limited range of validity of the frequency definition compared to the Bayesian one. For example, Laplace used the Bayesian definition to estimate the most probable mass of Saturn, given orbital data from various astronomical observatories. However, according to the frequency definition, we are not permitted to use probability theory to tackle this problem, because the mass of Saturn is not

a random variable, but rather a constant. Therefore, it does not have a frequency distribution. This is also the case for hypotheses, so in frequentist probability theory we can not directly calculate the probability of a hypothesis. We can only indirectly infer the truth of a hypothesis, in contrast to Bayesian probability theory, where the probability of any hypothesis can be assessed directly.

Let us illustrate some of the problems associated with frequentist probability theory on the basis of the standard example of a coin toss experiment, where the probability of heads is widely agreed to be $1/2$. Two justifications are commonly given:

1. The symmetry of exchangeability argument:

$$\text{probability} = \frac{\text{number of favorable cases}}{\text{number of possibilities}}. \quad (5.1)$$

Here, it is assumed that all possibilities are equally likely. This is the case for an ideal coin (or an ideal dice), but in reality, this depends very much on assumptions on the physical characteristics of the coin, and of the nature of the physical forces at work.

2. The frequency argument: probability = relative frequency obtained in a ‘very long’ sequence of tosses, assumed to be performed in an identical matter, and physically independently of each other.

It is here that the subjectivity of the frequentist definition of probability becomes apparent. Both of the above arguments are in a sense subjective, since they require judgements about the nature of the coin and the tossing procedure. It is not entirely clear what is meant by ‘equally likely events’, ‘identical measurements’ and ‘independence’. In addition, the frequency argument has certain special difficulties, in that it involves the hypothetical notion of a very long sequence of identical tosses. But what if we want to know the probability of heads for a non-ideal coin? Should we perform the experiment in reality, and how do we construct a sequence of *identical* tosses? These problems become even more apparent when we want to assess the probability that it will rain tomorrow. Or, suppose Colombia plays against Brazil in soccer, what is the probability that Colombia wins? What is the probability that Colombia wins if it rains tomorrow? What is the probability that the next space shuttle launched will explode? It is clear that in these

cases we can not perform a long sequence of identical experiments, but a frequency interpretation is usually constructed, by considering the system under study as being part of a hypothetical ensemble of identical and independent systems. For example, one might look at the frequency of past shuttle launches that have exploded. However, again it is not entirely clear, and at least dependent on a great deal of subjectivity from the part of the statistician, how the characteristics of this reference set of ‘identical’ experiments correspond to the real physical experiment.

In contrast, in Bayesian probability theory only the data that were actually measured, combined with relevant prior information, is considered. Hypothetical data sets, taken from an ensemble of systems, play no role. The system for assigning probabilities in Bayesian probability theory is outlined below.

Thus, frequentist mathematicians soon realized that the frequency definition of probability did not permit to solve most real-life scientific problems, and a new subject was invented: statistics. For example, to estimate the mass of Saturn, one has to relate the mass to the data through some function called a *statistic*. Then, since the data are subject to ‘random’ noise, the statistic becomes the variable to which the rules of probability theory can be applied. The question now arises how to choose the statistic. The frequentist approach does not yield a natural, or objective, way of doing this, and many tests and procedures were invented to cope with this problem.

In 1946, Richard Cox [77] tried to get away from the controversy of the Bayesian versus frequentist point of view. He constructed a set of rules that are necessary for logical and consistent reasoning. He started by considering how one might express one’s relative beliefs in the truth of a proposition. Cox assumed that, to obtain a transitive property of this degree of belief, it would be the most straightforward to assign a real number to a state of belief. He also imposed the constraint that if there are several ways of using the same information, then one should always arrive at the same conclusion. By using Boolean logic, he found that this consistency could only be ensured if the real numbers attached to degrees of belief, could be mapped onto another set of *positive* real numbers, which obeyed the usual rules of probability theory:

$$P(X|I) + P(\bar{X}|I) = 1, \quad \text{sum rule,} \quad (5.2)$$

and

$$P(X, Y|I) = P(X|Y, I) \times P(Y|I), \quad \text{product rule.} \quad (5.3)$$

Here, X and Y denote two propositions, while \bar{X} denotes the proposition that X is false. I is the proposition expressing any additional information that is available. This additional, or *prior*, information is very important in Bayesian probability theory. It expresses that there is no such thing as an absolute probability¹. For example, the probability that we assign to the proposition ‘it will rain tomorrow’ depends e.g. on whether there are dark clouds or a clear sky today. The explicit subjectivity of Bayesian probability theory is a reason for a lot of criticism. However, the subjective nature of probability is really not an issue, since we can require that observers with the same amount of information should come to the same conclusion. Moreover, it is possible to assume as little as possible, i.e. to select prior information that is as uninformative as possible (see Section 5.1.5). The frequentist approach gives the impression to be a more objective theory, but in reality it just makes life more complicated.

In summary, the main advantages of the Bayesian approach are the following (some concepts are clarified later on):

1. all variables, including model parameters, can be assigned a probability distribution.
2. it provides a simple approach for answering any probabilistic question, for a given state of information.
3. it calculates the probability of a hypothesis H_i directly: $p(H_i|D, I)$, where D are the data.
4. it incorporates relevant prior information through Bayes’ theorem. This is especially useful when there are little data or when the data are very noisy.

¹This inherent subjectivity of probabilities raises a problem in the probabilistic (Kopenhagen) interpretation of quantum mechanics, where probabilities of physical events are usually assumed to be objective. However, this ambiguity can be solved in the Bayesian interpretation of quantum mechanics, rendering all probabilities Bayesian, while retaining the objectivity of the wave function [78].

5. it provides a straightforward way to incorporate the effect of systematic errors on the data via nuisance parameters, which can be eliminated through marginalization.
6. it automatically prevents overfitting of a model by incorporating Ockham's razor.

More information on Bayesian probability theory can be found under Refs. [79], [80] and [81].

5.1.2 Bayes' theorem

The sum and product rule, Equations (5.2) and (5.3), respectively, form the basic algebra of probability theory. From this, one can derive many results. For example, interchanging X and Y in (5.3), we get

$$P(Y, X|I) = P(Y|X, I) \times P(X|I). \quad (5.4)$$

Since the probability of both X and Y being true, is logically the same as the probability that Y and X is true, (5.3) should equal (5.4):

$$P(X|Y, I) \times P(Y|I) = P(Y|X, I) \times P(X|I).$$

Hence, we arrive at *Bayes' theorem* (or Bayes' rule).

Theorem 1 (Bayes). *For the propositions X and Y , and a given prior proposition I :*

$$P(Y|X, I) = \frac{P(X|Y, I) \times P(Y|I)}{P(X|I)}.$$

Thus, Bayes' theorem allows to transpose the conditional probability $P(X|Y, I)$ to $P(Y|X, I)$. In the current context, we will often work with a set of measured data, represented by the vector \mathbf{x} , and a set of parameters characterizing a certain physical system, represented by the vector $\boldsymbol{\theta}$. Instead of dealing with the probability of a proposition $P(X)$, we will rather turn to the (multivariate) Probability Density Function (PDF) of a set of variables $p(\mathbf{x})$. In this notation, Bayes' theorem can be written as

$$p(\boldsymbol{\theta}|\mathbf{x}, I) = \frac{p(\mathbf{x}|\boldsymbol{\theta}, I)p(\boldsymbol{\theta}|I)}{p(\mathbf{x}|I)}, \quad (5.5)$$

where I still represents any additional information at hand. This includes for example additional information on the physical system under study, i.e. the physical model that is used.

The factor $p(\mathbf{x}|\boldsymbol{\theta}, I)$ in the nominator of the right hand side of (5.5) is called the *likelihood* of the parameter vector $\boldsymbol{\theta}$. To emphasize that it is a function of $\boldsymbol{\theta}$, it is sometimes written as $L(\boldsymbol{\theta})$. The likelihood is the probability (density) to observe the data vector \mathbf{x} , *given* a set of model parameters $\boldsymbol{\theta}$. It typically involves a difference between the observed data \mathbf{x} , and the data calculated from the given set of parameters $\boldsymbol{\theta}$, via the assumed physical model, called the *forward model* or *data descriptive model*. Thus, the likelihood describes the data misfit.

The factor $p(\boldsymbol{\theta}|I)$ is called the *prior* probability (density) of the parameter vector $\boldsymbol{\theta}$, and it embodies everything we are willing to assume about the model parameters, *before* gathering the data. The latter is imperative, and an often made mistake is to use the data for constructing a prior PDF. This does not prevent, however, to employ information on data that was gathered in a *previous* experiment for defining a prior.

The quantity on the left hand side of (5.5) is named the *posterior* PDF. It is the PDF of interest, since it gives the probability (density) of the parameters of interest $\boldsymbol{\theta}$ of the physical model (e.g. the mass of Saturn), *given* the observed data (e.g. orbital data of Saturn). This problem of determining the probability for the underlying parameters of a physical model, given a relevant set of measurements, is precisely the inverse problem we mentioned earlier in Section 5.1.1.

The factor $p(\mathbf{x}|I)$ in the denominator of (5.5) is named the *evidence*. It does not depend on the parameters $\boldsymbol{\theta}$, and it is often ignored since it merely normalizes the posterior. However, this factor is important during the task of model selection, as we will demonstrate shortly.

5.1.3 Marginalization

In the case where there is more than one model parameter, Bayes' theorem yields a joint PDF for the parameter vector $\boldsymbol{\theta}$:

$$p(\boldsymbol{\theta}|\mathbf{x}, I) = p(\theta_1, \dots, \theta_p|\mathbf{x}, I), \quad (5.6)$$

where p is the number of parameters. However, we typically want to make inferences about individual parameters $\theta_i, i = 1, \dots, p$. In fact, the posterior often depends also on parameters we are not even interested in, but that necessarily enter the data descriptive model. These are called *nuisance parameters*. In order to obtain the PDF for an individual parameter θ_i , we have to *marginalize* the joint PDF, i.e.

integrate out the other parameters:

$$\begin{aligned} p(\theta_i|\mathbf{x}, I) &= \int p(\boldsymbol{\theta}|\mathbf{x}, I) d\theta_1 \dots d\theta_{i-1} d\theta_{i+1} \dots d\theta_p \\ &\sim \int \frac{p(\mathbf{x}|\boldsymbol{\theta}, I)p(\boldsymbol{\theta}|I)}{p(\mathbf{x})} d\theta_1 \dots d\theta_{i-1} d\theta_{i+1} \dots d\theta_p. \end{aligned} \quad (5.7)$$

The integral is over the complete parameter space spanned by the parameters $\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p$, and it may also be a sum in the case of discrete parameters. In addition, in order to characterize the posterior PDF, we might want to calculate its moments. To do this, we also need to integrate the posterior over the model parameters.

The integrations can be very difficult to perform, and we will come back to this problem in Section 5.2.

5.1.4 Bayesian model selection and Ockham's razor

Since Bayesian probability theory allows to ask for the probability of any event or quantity, one can also evaluate the probability of a certain physical model to be true, given a set of data. This probability can then be compared to the probability of other models, and a most probable model can be identified. This is the concept of *Bayesian model selection*, which is one of the great triumphs of Bayesian probability theory, that distinguishes it from frequentist probability theory. Model selection has only recently been developed in depth.

In practice, the posterior probability for a certain model \mathcal{H}_i to be correct, after performing an experiment that resulted in a set of data \mathbf{x} , is through Bayes' theorem given by the following proportionality:

$$p(\mathcal{H}_i|\mathbf{x}, I) \sim p(\mathbf{x}|\mathcal{H}_i, I)p(\mathcal{H}_i|I).$$

Here, I denotes any prior information we might have about the truth of the model \mathcal{H}_i . In the factor $p(\mathbf{x}|\mathcal{H}_i, I)$, we recognize the evidence (for the model \mathcal{H}_i), defined in the previous subsection, where the information on the model \mathcal{H}_i has been taken explicitly out of the rest of the prior information I . It is the probability that the model \mathcal{H}_i is able to realize the measured data set.

We now give a simple argument showing that Bayesian probability theory automatically incorporates Ockham's razor, penalizing too complex models. Our argument is based on a reasoning by MacKay [82].

We can write the evidence for the model \mathcal{H}_i as follows:

$$p(\mathbf{x}|\mathcal{H}_i, I) = \int p(\mathbf{x}|\boldsymbol{\theta}, \mathcal{H}_i, I)p(\boldsymbol{\theta}|\mathcal{H}_i, I)d\boldsymbol{\theta}, \quad (5.8)$$

with $\boldsymbol{\theta}$ the model parameters. For many problems the posterior $p(\boldsymbol{\theta}|\mathbf{x}, \mathcal{H}_i, I) \sim p(\mathbf{x}|\boldsymbol{\theta}, \mathcal{H}_i, I)p(\boldsymbol{\theta})$ has a strong peak at the most probable parameter vector $\boldsymbol{\theta}_{\text{MP}}$ (Figure 5.1). Then the evidence can be approximated by the height of the peak of the integrand $p(\mathbf{x}|\boldsymbol{\theta}, \mathcal{H}_i, I)p(\boldsymbol{\theta}|\mathcal{H}_i, I)$ times its approximate width, $\Delta\boldsymbol{\theta}$:

$$\underbrace{p(\mathbf{x}|\mathcal{H}_i, I)}_{\text{evidence}} \approx \underbrace{p(\mathbf{x}|\boldsymbol{\theta}_{\text{MP}}, \mathcal{H}_i, I)}_{\text{best fit likelihood}} \underbrace{p(\boldsymbol{\theta}|\mathcal{H}_i, I) \times \Delta\boldsymbol{\theta}}_{\text{Ockham factor}}.$$

$\Delta\boldsymbol{\theta}$ is a measure for the posterior uncertainty in $\boldsymbol{\theta}$. Imagine now that the prior $p(\boldsymbol{\theta}|\mathcal{H}_i, I)$ is uniform on some large interval $\Delta^0\boldsymbol{\theta}$. The latter represents the range of values for $\boldsymbol{\theta}$ that are thought to be possible before the data arrive. Hence, $p(\boldsymbol{\theta}|\mathcal{H}_i, I) = \frac{1}{\Delta^0\boldsymbol{\theta}}$, and we define:

$$\text{Ockham factor} \equiv \frac{\Delta\boldsymbol{\theta}}{\Delta^0\boldsymbol{\theta}}, \quad (5.9)$$

i.e. the ratio of the posterior accessible volume of the parameter space to the prior accessible volume. It is the factor by which the hypothesis space of the model collapses when the data arrive. Now, the more parameters are included in the model, describing the system under study, the smaller the Ockham factor will become, thereby decreasing the evidence and consequently the posterior probability for the model. Thus, although a complicated model, containing many parameters, may result in a higher best fit likelihood, the Ockham factor penalizes the model for its complexity. This way, possibly a less complicated model will be favored, as long as it explains the data reasonably well. This embodies the common effect of overfitting of a data set: a very complicated model can easily accommodate all data points very accurately, minimizing the data misfit, but intuitively we are more inclined to prefer a simpler model that still explains the data well. A simple example occurs in polynomial fitting of a data set, where a polynomial of high degree can fit the data very well. However, in doing so it oscillates wildly between the data points, and a polynomial of a lower degree will exhibit a more physically plausible behaviour. In Bayesian probability theory, the model that achieves the largest evidence is determined by a trade-off between maximizing the Ockham factor, and minimizing

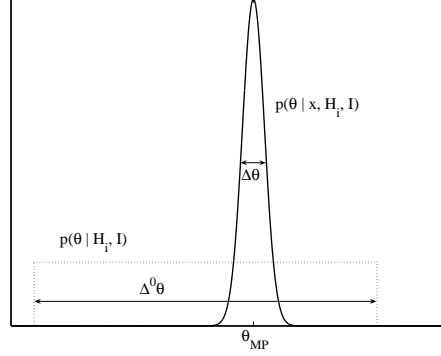


Figure 5.1: Illustration of the quantities that determine the Ockham factor for a model \mathcal{H}_i with a single parameter θ . The prior distribution for the parameter is uniform (dotted line) and has a width $\Delta^0 \theta$. The posterior (solid line) has a single maximum at θ_{MP} with characteristic width $\Delta \theta$.

the data misfit. Thus, Bayesian probability theory embodies a natural principle that rules out too complicated models.

Incidentally, the term ‘Ockham factor’ originates from the work of the fourteenth century English philosopher William of Ockham, to whom is attributed the following principle of parsimony in explanation and theory building:

Numquam ponenda est pluralitas sine necessitate.

This is usually understood as:

When multiple competing theories fit the data equally well, the one that introduces the fewest assumptions and parameters, should be selected.

5.1.5 Prior PDF selection

The subject of this subsection is how to quantify our prior knowledge into a suitable probability distribution. We will also demonstrate that, as the amount of measured data increases, the influence of the prior on the posterior decreases gradually. Hence, when a large amount of data is available, the choice of prior does not matter so much anymore. On the other hand, in cases where we have a limited data set, or when

the data quality is poor, the prior will be all the more important. Especially in such instances, we have to be cautious not to encode more information into the prior PDF than is in fact supported by the available knowledge. In addition, care has to be taken to construct a prior that does not lead to an *improper* posterior, i.e. one that can not be normalized because its integral over the parameter space diverges. We start with the case where we want to assume the least possible for the construction of a prior.

5.1.5.1 Uninformative prior distributions

Our desire here is to construct a prior distribution that plays a minimal role in the posterior distribution. The rationale for using uninformative prior distributions is ‘to let the data speak for themselves’, so that our inferences are the least possible affected by information external to the data at hand. Many principles exist to construct uninformative priors, all leading to (slightly) different results. But for two cases all principles seem to agree: *location parameters* and *scale parameters*.

A location parameter x_0 is a quantity that enters a probability distribution $p_{x_0}(x)$ like $p(x - x_0)$. An example is the mean value of a normal distribution. If x_0 is a location parameter for the likelihood in the problem under study, that is, $p(x - x_0|x_0)$ depends only on $x - x_0$, then it is reasonable that an uninformative prior distribution for x_0 would yield a posterior that still only depends on $x - x_0$, i.e. $p(x - x_0|x)$. Using Bayes’ theorem, we find $p(x - x_0|x) \sim p(x_0)p(x - x_0|x_0)$, implying that the uninformative density $p(x_0)$ is a uniform distribution.

A scale parameter s enters a probability distribution $p_s(x)$ like $p(x/s)$. Suppose s is a scale parameter for the likelihood in the problem under study: $p(x/s|s)$, or, by transformation of variables:

$$p(x|s) = \frac{1}{s}p(x/s|s). \quad (5.10)$$

An example of a scale parameter for such a likelihood is the standard deviation for a normal distribution. Again, it is reasonable that an uninformative prior would result in a posterior that has s as a scale parameter as well. Transformation of variables also gives

$$p(s|x) = \frac{x}{s^2}p(x/s|x). \quad (5.11)$$

Since $p(x/s|s) \equiv p(x/s|x)$, (5.10) and (5.11) yield

$$p(s|x) = \frac{x}{s}p(x|s),$$

such that an uninformative prior for s is $p(s) \sim 1/s$.

5.1.5.2 Conjugate prior distributions

A prior distribution is often chosen based on the property that the posterior follows the same parametric form as the prior, a feature called *conjugacy*. For example, if x is distributed according to an univariate binomial with parameter θ , then its distribution is of the form

$$p(x|\theta) \sim \theta^a(1 - \theta)^b.$$

If the prior is of the same form, namely:

$$p(\theta) \sim \theta^{\alpha-1}(1 - \theta)^{\beta-1}, \quad (5.12)$$

then the posterior will also be of this form. Incidentally, the distribution in (5.12) is a beta distribution with parameters α and β . The parameters α and β are called *hyperparameters*. In turn, these parameters can be assigned a prior distribution (*hyperprior*), possibly parameterized by *hyperhyperparameters*, and so on. Thus, one creates a so-called *hierarchical* probabilistic model, possibly consisting of several levels.

Up to now, the main motivation for choosing a conjugate prior is mathematical convenience. Of course, if information is available that contradicts the conjugate parametric family, it may be necessary to use a more realistic prior, albeit a less convenient one. However, recent research suggests that conjugate priors are a special case of a larger and more general class of priors known as *entropic priors*, which maximize ignorance [83]. So, in fact conjugacy seems to be a good choice after all, if the aim is to assume the least possible.

5.1.5.3 Encoding expert knowledge

Finally, if prior information is available, and we are willing to trust it, we can quantify it in the prior distribution. This kind of prior knowledge is often called *expert knowledge*, as it is assumed to originate from an 'expert's opinion. Relevant prior knowledge can be based on data from previous similar experiments, and might concern information on the value of the mean or variance of the distribution of a parameter. Often, a normal distribution is taken for a prior, since many physical quantities are approximately normally distributed, by virtue of the Central Limit Theorem.

5.1.6 Influence of the prior on the posterior

We will here demonstrate, by a simple example from Bayesian estimation, that the influence of the prior distribution on the posterior depends on the amount of data available. As one might expect, this influence decreases as more data are gathered.

We consider the estimation of the mean θ of a normal distribution with a known variance σ^2 . Suppose at first there is only a single observation x , assumed to have been sampled from

$$p(x|\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\theta)^2}.$$

This is the likelihood as a function of the parameter θ , which is an exponential of a quadratic form in θ . The family of conjugate priors can be written as

$$p(\theta) = e^{a\theta^2 + b\theta + c},$$

which we reparameterize as

$$p(\theta) \sim e^{-\frac{1}{2\tau_0^2}(\theta - \mu_0)^2};$$

that is, $\theta \sim N(\mu_0, \tau_0^2)$, with hyperparameters μ_0 and τ_0^2 . We assume the hyperparameters to be known.

The posterior distribution for θ is the exponential of a quadratic form, and thus normal. After some algebra, we find:

$$p(\theta|x) \sim \exp\left(-\frac{1}{2}\left[\frac{(x-\theta)^2}{\sigma^2} + \frac{(\theta - \mu_0)^2}{\tau_0^2}\right]\right),$$

or

$$p(\theta|x) \sim e^{-\frac{1}{2\tau_1^2}(\theta - \mu_1)^2};$$

that is, $\theta|x \sim N(\mu_1, \tau_1^2)$, where

$$\mu_1 \equiv \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}x}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}} \quad \text{and} \quad \frac{1}{\tau_1^2} \equiv \frac{1}{\tau_0^2} + \frac{1}{\sigma^2}. \quad (5.13)$$

In manipulating normal distributions, the inverse variance plays a prominent role and is called the *precision*, which we will denote by the symbol γ . Thus, from (5.13):

$$\gamma_1 = \gamma_0 + \gamma_D,$$

where $\gamma_D = \frac{1}{\sigma^2}$ denotes the data precision. We can conclude that for normally distributed data and a normal prior, *the posterior precision equals the prior precision plus the data precision*. On the other hand, *the posterior mean μ_1 can be seen as a weighted average of the prior mean and the observed value, x , with weights proportional to the precisions*. The posterior mean is thus a compromise between the prior mean and the observed value.

As $\gamma_0 \rightarrow \infty$, the prior becomes infinitely more precise than the data, and so the posterior and prior distributions are identical and concentrated at the value μ_0 . As, on the other hand, $\gamma_D \rightarrow \infty$, the data become exact, and the posterior is concentrated at the observed value, x . If $x = \mu_0$, the prior and data means coincide, and the posterior mean must also lie at this point.

In the case of the more realistic situation where there are multiple observations, the preceding results can easily be extended. Assume that a set of n independent observations $\mathbf{x} = (x_1, \dots, x_n)$ has been sampled from the same normal distribution with unknown mean θ , and known variance σ^2 . The posterior is

$$\begin{aligned} p(\theta|\mathbf{x}) &\sim p(\theta)p(\mathbf{x}|\theta) \\ &= p(\theta) \prod_{i=1}^n p(x_i|\theta) \\ &\sim \exp\left(-\frac{1}{2\tau_0^2}(\theta - \mu_0)^2\right) \prod_{i=1}^n \exp\left(-\frac{1}{2\sigma^2}(x_i - \theta)^2\right) \\ &\sim \exp\left(-\frac{1}{2}\left[\frac{1}{\tau_0^2}(\theta - \mu_0)^2 + \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \theta)^2\right]\right). \end{aligned}$$

Algebraic simplification shows that the posterior depends on \mathbf{x} only through the sample mean $\bar{x} \equiv \frac{1}{n} \sum_i x_i$. In fact, since

$$\bar{x}|\theta, \sigma^2 \sim N(\theta, \sigma^2/n),$$

the results derived for the single normal observation apply immediately, treating \bar{x} as the single observation:

$$p(\theta|x_1, \dots, x_n) = p(\theta|\bar{x}) = N(\theta|\mu_n, \tau_n^2),$$

where

$$\mu_n \equiv \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{x}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}} \quad \text{and} \quad \frac{1}{\tau_n^2} = \frac{1}{\tau_0^2} + \frac{n}{\sigma^2}.$$

From this, we can conclude that, if n is large, the posterior is largely determined by the data variance σ^2 and the sample average \bar{x} . This means that the choice of prior will gradually play a smaller role as more data arrive. $\frac{n}{\sigma^2}$ is the data precision, hence, the more data arrive, the higher the precision of the posterior, according to our expectations. If now $\tau_0^2 = \sigma^2$, then the prior distribution has the same weight as one extra observation with the value μ_0 . As the prior information becomes more diffuse, i.e. $\tau_0 \rightarrow \infty$, with n fixed, or as $n \rightarrow \infty$ with τ_0^2 fixed, we have:

$$p(\theta|\mathbf{x}) \approx N(\theta|\bar{x}, \sigma^2/n).$$

This is in practice a good approximation for the posterior whenever prior beliefs are relatively diffuse over the range of θ where the likelihood is substantial.

5.2 Bayesian computational methods

In order to summarize our inferences concerning the physical model under study, it is often convenient to estimate moments of the marginal posterior distribution of the respective model parameters. In doing so, we have to calculate integrals over the joint posterior distribution. For example, to calculate the posterior mean of the parameter θ_i in a model with parameters $\theta_1, \dots, \theta_p$, we have to perform the integration:

$$\int \theta_i p(\boldsymbol{\theta}|\mathbf{x}, I) d\theta_1 \dots d\theta_p.$$

As mentioned before, such integrations can be very difficult to carry out, because the joint posterior might in fact be a very complicated (intractable) function of the parameters. In addition, many problems in Bayesian inference have a lot of parameters, possibly hundreds. At this point, even numerical integration methods become very cumbersome. In the analysis described in Chapter 6, we will therefore, when appropriate, resort to *stochastic* integration methods, and Markov Chain Monte Carlo (MCMC) simulation in particular. In this section, we will summarize the main results of MCMC theory, but we will almost always omit the proofs of the stated theorems. These can be found e.g. under Ref. [84] and references therein. Further information on MCMC methods is available under Refs. [85], [86] and [81].

5.2.1 Some historic notes

We here give a short historic background on Monte Carlo methods, and Markov Chain Monte Carlo. Many of the concepts mentioned are explained later on.

The origins of the Monte Carlo method lie in the Manhattan project, the United States government's plan to build the first atomic bomb during World War II. There, Nicholas Metropolis, Richard Feynman and John von Neumann became interested in the possibilities of fast electronic calculators. Later, also at Los Alamos, Stanislaw Ulam and John von Neumann thought of the idea of performing computations on physical systems via probabilistic simulation. Apparently Metropolis proposed the catchy name 'Monte Carlo methods' [87].

In their famous paper of 1953, Metropolis *et al.* proposed a method for the computation of a $2N$ -dimensional integral that is a function of the energy E of a physical system [88]. To do this, N particles are put randomly at points on a square, the energy E of this configuration is calculated, the configuration is weighted by the Boltzmann factor $\exp(-E/kT)$, and Monte Carlo integration is performed. Here, k is Boltzmann's constant and T is the temperature in the system. Statisticians would later call $\exp(-E/kT)$ the *objective function*. Metropolis and co-workers thus described a method that is known today as *simulated annealing*, in which each particle on the square is moved according to a random (uniform) perturbation, forming a new configuration. This new configuration is always accepted if the change in energy ΔE is negative, the underlying principle here being that a physical system tends for the lowest energy configuration. However, if $\Delta E > 0$, the new configuration is accepted with probability $\exp(-\Delta E/kT)$, otherwise the previous one is retained. Simulated annealing is a popular method for solving (possibly high-dimensional) optimization problems. The purpose is to maximize a function $h(\boldsymbol{\theta})$ of the system variables. By defining h as the negative 'energy' of the system, simulated annealing moves on the energy landscape, trying to find an absolute minimum of the energy. In allowing for a nonzero acceptance probability for a configuration with a *higher* energy than the previous one, simulated annealing avoids getting trapped in local energy minima; a very important feature of the method.

So far, the Metropolis method seemingly had its use specifically in solving problems in statistical mechanics. However, in 1970 the method

was generalized and improved by Hastings [89], who viewed the Metropolis algorithm chiefly as a way to sample from high-dimensional probability distributions, which is precisely its primary modern use. Hastings noted that the Metropolis method involved the transition matrix of a Markov chain. He presented the target distribution to sample from in terms of the invariant distribution π of the Markov chain. Thus, the Metropolis-Hastings (M-H) algorithm was born, to become member of a larger family of what would later be known as Markov Chain Monte Carlo (MCMC) methods.

In the early 1990s, another popular MCMC method, known as *Gibbs sampling*, was proposed, mainly due to the work of Gelfand and Smith [90], who built on the paper of Geman and Geman [91]. The Gibbs sampler is a related simulation algorithm that is especially useful for sampling multivariate distributions, particularly when the univariate conditional densities are known, or are easy to sample from. From this moment, MCMC methods became, gradually, widely used in mainstream statistics. Later, it was shown that the Gibbs sampling algorithm is in fact a special case of M-H ([92], and see also below). In 1994, Tierney wrote an influential article summarizing the history and theory of the M-H algorithm [93], and showed how it (and the Gibbs sampler) could be employed to sample from any (intractable) posterior distribution, often arising in Bayesian inference.

5.2.2 Markov chains

We start our introduction into the field of MCMC methods with a brief discussion of the fundamental notions of Markov chains that are needed to study MCMC algorithms. For a more thorough introduction to Markov chains, we refer to the work of Meyn and Tweedie [94].

We first note that we do not deal here with Markov models in continuous time (also called ‘Markov processes’), since the nature of simulation automatically leads us to consider only discrete-time stochastic processes.

Definition 1. A Markov chain is a sequence of random variables $X^{(0)}, X^{(1)}, X^{(2)}, \dots$ where, for all $t \in \mathbb{N}$, the distribution of $X^{(t+1)}$ depends only on the value $x^{(t)}$ of $X^{(t)}$:

$$p(X^{(t+1)} | x^{(0)}, x^{(1)}, \dots, x^{(t)}) = p(X^{(t+1)} | x^{(t)}).$$

We will call t the *time parameter*, and we denote the conditional probability density $p(X^{(t+1)}|x^{(t)})$ by $K(X^{(t)}, X^{(t+1)})$, the *transition kernel* of the chain. This means that

$$X^{(t+1)} \sim K(X^{(t)}, X^{(t+1)}).$$

For two arbitrary states $x, y \in \mathcal{X}$, we will sometimes also use the notation $K(x, y)$. The $X^{(i)}$ have a common range, the *state space* \mathcal{X} of the Markov chain. If we denote by $\mathcal{B}(\mathcal{X})$ the Borel σ -algebra on \mathcal{X} , we also have that for any subset $A \in \mathcal{B}(\mathcal{X})$:

$$P(X \in A|x) = \int_A K(x, y)dy.$$

The notation $K^n(x, y)$ is also sometimes used, which represents the kernel for $n \in \mathbb{N}^*$ transitions. If we define $K^1(x, y) \equiv K(x, y)$, then for $n > 1$:

$$K^n(x, y) = \int_{\mathcal{X}} K^{n-1}(x', y)K(x, dx')$$

We have applied a slight abuse of notation here, since by $K(x, dx')$ we actually mean $K(x, x')$ with $dx' \equiv x' - x$. Finally, for every $A \in \mathcal{B}(\mathcal{X})$, the kernel $X(x, A)$ denotes a transition $X^{(t)} \rightarrow X^{(t+1)}$ from the state x to the subset A , i.e. $X^{(t+1)} \in A$.

Formally, the following definition can be given for the transition kernel.

Definition 2. *A transition kernel is a function K defined on $\mathcal{X} \times \mathcal{B}(\mathcal{X})$ such that*

- (i) $\forall x \in \mathcal{X}$, $K(x, \cdot)$ is a probability measure,
- (ii) $\forall A \in \mathcal{B}(\mathcal{X})$, $K(\cdot, A)$ is measurable.

By virtue of the marginalization principle, we also have

$$p(X^{(t+1)}) = \int_{\mathcal{X}} p(X^{(t)})K(X^{(t)}, dx).$$

A Markov chain can be specified entirely by giving the distribution for $X^{(0)}$, as well as the conditional distributions for $X^{(t+1)}$ given $X^{(t)}$, $\forall t$, i.e. the complete transition kernel. If \mathcal{X} is discrete, the transition kernel becomes a transition matrix K with elements

$$P_{xy} \equiv P(X^{(t+1)} = y | X^{(t)} = x), \quad x, y \in \mathcal{X}.$$

Definition 3. A Markov chain $(X^{(i)})$ is said to be (time-)homogeneous if its transition kernel does not depend on the time parameter.

In the following, we will only work with homogeneous Markov chains. We are particularly interested in Markov chains whose distribution converges to a certain limiting distribution, which is the target distribution of interest, which we want to sample from. Therefore, we want to reduce the sensitivity of the chain to the initial conditions $X^{(0)}$.

Definition 4. A definite probability density π is said to be stationary (or invariant w.r.t. a Markov chain $(X^{(i)})$) if and only if

$$\pi(X^{(t+1)}) = \int_{\mathcal{X}} \pi(X^{(t)})K(X^{(t)}, dx), \quad \forall t.$$

This means that, once the distribution π has been reached, it will always stay the distribution of the chain, and the chain is said to be stationary. When there exists an invariant probability density for a Markov chain, the chain is called *positive*. We will later see another definition for positivity.

For a chain to converge to a stationary distribution, it needs to satisfy three important properties.

1. The chain needs to be *irreducible*, meaning that the chain can reach any state with nonzero probability, in a finite number of steps, independent of the starting conditions.
2. The chain must be *aperiodic*, so that it does not oscillate between different sets of states in a periodic fashion.
3. The chain should be *positive recurrent*.

We now formalize these ideas.

Definition 5. Given a measure φ , the Markov chain $(X^{(i)})$ with transition kernel $K(x, y)$ is φ -irreducible if, for every $A \in \mathcal{B}(\mathcal{X})$ with $\varphi(A) > 0$, there exists an $n \in \mathbb{N}^*$ such that $K^n(x, A) > 0$ for all $x \in \mathcal{X}$.

We give the following theorem without proof.

Theorem 2. If a Markov chain is irreducible and it allows for a stationary distribution π , then π is unique, up to a multiplicative factor.

Definition 6. A set C is small if there exist $m \in \mathbb{N}^*$ and a nonzero measure ν_m such that

$$K^m(x, A) \geq \nu_m(A), \quad \forall x \in C, \forall A \in \mathcal{B}(\mathcal{X}).$$

Definition 7. A φ -irreducible Markov chain $(X^{(i)})$ has a cycle length d if there exists a small set C , an associated integer m , and a probability distribution ν_m such that $d \in \mathbb{N}^*$ is the greatest common denominator of

$$\{m \geq 1; \exists \delta_m > 0 : C \text{ is small for } \nu_m \geq \delta_m \nu_m\}. \quad (5.14)$$

The period of the chain is defined as the largest integer d satisfying (5.14), and the chain is aperiodic if $d = 1$.

Although irreducibility already guarantees that every set $A \in \mathcal{B}(\mathcal{X})$ will be visited by the Markov chain, for practical MCMC purposes this property is too weak to ensure that the trajectory of the chain will enter A often enough, and we need to require *recurrence* of the chain. We denote by η_A the number of visits of the chain to the set A :

$$\eta_A \equiv \sum_{i=0}^{\infty} \mathbb{I}_A(X^{(i)}),$$

where \mathbb{I}_A is the indicator function on A , satisfying

$$\begin{cases} \mathbb{I}_A(X^{(i)}) = 1 & \text{if } X^{(i)} \in A, \\ \mathbb{I}_A(X^{(i)}) = 0 & \text{otherwise.} \end{cases}$$

The notation $\mathbb{E}(\eta_A)$ refers to the expected number of visits to A .

Definition 8. A Markov chain $(X^{(i)})$ is recurrent if

- (i) there exists a measure φ such that the chain is φ -irreducible, and
- (ii) for every $A \in \mathcal{B}(\mathcal{X})$ such that $\varphi(A) > 0$, $\mathbb{E}(\eta_A) = \infty$.

Let τ_{xx} be the time of first return to state x , i.e.:

$$\tau_{xx} = \min\{t > 0 : X^{(t)} = x | X^{(0)} = x\}.$$

Definition 9. A Markov chain $(X^{(i)})$ is positive (recurrent) if $\mathbb{E}(\tau_{xx}) < \infty$, for every state x .

Incidentally, we also have the following proposition (without proof).

Proposition 1. *If a Markov chain is positive, it is recurrent (hence positive recurrent).*

Although not strictly necessary, the stability conditions for a Markov chain can even be increased, if one not only requires an infinite average number of visits to any subset A , but also an infinite number of visits for every path of the Markov chain. We denote by $P(\eta_A = \infty)$ the probability of visiting A an infinite number of times.

Definition 10. *A set A is Harris recurrent if $P(\eta_A = \infty) = 1$. The Markov chain $(X^{(i)})$ is Harris recurrent if there exists a measure φ such that $(X^{(i)})$ is φ -irreducible and for every set A with $\varphi(A) > 0$, A is Harris recurrent.*

Definition 11. *A Markov chain $(X^{(i)})$ is ergodic if it is aperiodic and positive Harris recurrent.*

It turns out that for an ergodic Markov chain, the (unique) stationary distribution π (which will be our target distribution) is also the limiting distribution, i.e. the chain's distribution (the kernel) *converges* to π . Importantly, this is true regardless of the initial conditions of the chain.

Theorem 3. *If the Markov chain $(X^{(i)})$ is ergodic, then*

$$\lim_{n \rightarrow \infty} \left\| \int K^n(x, \cdot) \mu(dx) - \pi \right\| = 0,$$

for every initial distribution μ .

Inferences drawn from running an MCMC simulation are usually summarized in terms of *ergodic averages* of the form

$$\bar{f}_n \equiv \frac{1}{n} \sum_{i=1}^n f(X^{(i)}),$$

f being a function of interest of the chain states. For instance, if $f(x) \equiv x$, then \bar{f}_n equals the sample average of the chain states. Now, let $L^1(\pi)$ represent the class of functions that are integrable over a distribution π . The following crucial theorem is the Law of Large Numbers for Markov chains, and is commonly called the *Ergodic Theorem*.

Theorem 4 (Ergodic). *If the Markov chain $(X^{(i)})$ is ergodic and $f \in L^1(\pi)$, then*

$$P \left(\bar{f}_n \rightarrow \int f(x) \pi(x) dx \right) = 1.$$

We will omit the proof. The Ergodic Theorem guarantees the convergence of the ergodic averages. For example, again if we take $f(x) \equiv x$, then \bar{f}_n converges surely to the mean of the stationary (target) distribution π , and similarly for other moments of the distribution.

We mention yet another stability property of a Markov chain, namely *reversibility*, which states that the direction of time does not matter.

Definition 12. *A stationary Markov chain $(X^{(i)})$ is reversible if the distribution of $X^{(t+1)}$ conditionally on $X^{(t+2)}$ is the same as the distribution of $X^{(t+1)}$ conditionally on $X^{(t)}$.*

In fact, reversibility can even be linked with the existence of a stationary distribution, if a condition stronger than in Definition 12 holds.

Definition 13. *A Markov chain with transition kernel K satisfies the detailed balance condition if there exists a function π satisfying*

$$K(y, x)\pi(y) = K(x, y)\pi(x),$$

for every (x, y) .

This means that the probability of being in the state x and moving to y is the same as the probability of being in y and moving back to x . When π is a probability density, this also implies that the chain is reversible.

Theorem 5. *If a Markov chain with transition kernel K satisfies the detailed balance condition with a probability density π , then:*

- (i) *the density π is the stationary density of the chain,*
- (ii) *the chain is reversible.*

We omit the proof. This theorem provides a sufficient (but not necessary) condition for a density π to be the stationary density of a Markov chain.

5.2.3 The principle of Monte Carlo integration

The principle of Monte Carlo integration is very simple. For simplicity, we work in one dimension, and we want to evaluate the following integral:

$$\int_{\mathcal{X}} h(x)f(x)dx. \tag{5.15}$$

To do this, we can view $f(x)$ as a (not necessarily normalized) probability density, so that (5.15) becomes the expected value of h under f : $\mathbb{E}_f(h)$. We now generate a (large) sample x_1, \dots, x_m from f . Then, we can approximate (5.15) by the empirical average

$$\bar{h}_m = \frac{1}{m} \sum_{j=1}^m h(x_j). \quad (5.16)$$

Indeed \bar{h}_m converges almost surely to (5.15) due to the Strong Law of Large Numbers. As a special case, by setting $h(x) = 1, x, x^2$, etc, we can approximate the moments of the distribution f , explaining the relevance of Monte Carlo integration in the current context.

5.2.4 Markov Chain Monte Carlo methods

In Bayesian statistics, Markov Chain Monte Carlo methods are used to sample from arbitrary marginal posterior distributions, by drawing samples from approximate distributions, and then correcting those draws to better approximate the target distribution. The samples are drawn sequentially, with the distribution of the draws depending on the last value drawn; hence the draws form a Markov chain. Markov chain simulation is used when it is not possible (or not computationally efficient) to sample directly from the posterior. In addition, Markov chain and other iterative simulation methods have many applications outside Bayesian statistics, such as in optimization, which we will not discuss here. The basic MCMC principle, to sample from arbitrary distributions, is to set up an ergodic Markov chain $(X^{(i)})$ whose limiting distribution is the target distribution of interest π . As explained before, by calculating ergodic averages, representing the Monte Carlo aspect of MCMC, we can approximate the moments of π , thus summarizing its characteristics. In practice, for an arbitrary starting value $x^{(0)}$, a chain $(X^{(i)})$ is generated using a transition kernel with stationary distribution π , which ensures the convergence in distribution of the chain to a random variable from π . Given that the chain is ergodic, the starting value $x^{(0)}$ is (in principle) unimportant.

Definition 14. A Markov Chain Monte Carlo (MCMC) method for the simulation of a distribution π is any method producing an ergodic Markov chain $(X^{(i)})$ whose stationary distribution is π .

5.2.4.1 The Metropolis-Hastings algorithm

Based on the above discussion of Markov chains, we will now derive the algorithm that will use to sample from our distributions of interest, namely the *Metropolis-Hastings* (M-H) algorithm.

We begin our discussion in the univariate case. In the M-H algorithm, a conditional density $q(y|x)$, which is easy to sample from, has to be chosen. The target density π must also be available up to some extent: a general requirement is that the ratio

$$\pi(y)/q(y|x)$$

is known up to a constant *independent* of x . This is crucial in Bayesian simulation, since we often only know the target posterior distribution up to a multiplicative factor. It is important to note that knowledge of the functional form of the target density π does *not* automatically allow to generate samples from π . Hence, for complicated distributions, or for a problem with a high dimensionality, we need MCMC to sample from these distributions, and to calculate moments.

The M-H algorithm associated with the objective (target) density π and the conditional density q produces a Markov chain $(X^{(i)})$ through the following transition.

Algorithm 1 (Metropolis-Hastings).

- 1: Given $x^{(t)}$
- 2: **repeat**
- 3: Generate $Y^{(t)} \sim q(y|x^{(t)})$.
- 4: Take

$$X^{(t+1)} = \begin{cases} Y^{(t)} & \text{with probability } \rho(x^{(t)}, Y^{(t)}), \\ x^{(t)} & \text{with probability } 1 - \rho(x^{(t)}, Y^{(t)}), \end{cases},$$

where

$$\rho(x, y) \equiv \min \left\{ \frac{\pi(y) q(x|y)}{\pi(x) q(y|x)}, 1 \right\}$$

- 5: **until** Satisfied

The distribution q is called the *proposal* (or *instrumental*) *distribution*, and the probability $\rho(x, y)$ the *Metropolis-Hastings acceptance probability*. Only in the symmetric case, where $q(x, y) \equiv q(y, x)$ does the acceptance depend solely on $\pi(y^{(t)})/\pi(x^{(t)})$, in which case the algorithm is simply called the *Metropolis algorithm*. In the general case, if the ratio $\pi(y^{(t)})/q(y^{(t)}, x^{(t)})$ is increased by the proposed value $y^{(t)}$, the proposal will always be accepted. Thus, the algorithm preferentially samples in regions of probability space where the probability mass of π is concentrated, while not wasting too much time in other regions. On the other hand, π might be multimodal, and in order not to get stuck in local maxima of π , from time to time also proposals that *decrease* the ratio $\pi(y^{(t)})/q(y^{(t)}, x^{(t)})$ are accepted. These are two major advantages of MCMC.

The transitions of the chain depend on q . Transitions should not be taken too small, because we want the support of the target distribution to be explored efficiently. On the other hand, neither should the transitions be taken too large, because then the chain is more likely to waste time in regions where the target distribution is small.

The samples that are generated by the M-H algorithm are not mutually independent. For one thing, there may be repeated occurrences of the same value. However, this poses no real problem, since we are merely interested in the convergence of empirical averages

$$\frac{1}{T} \sum_{i=1}^T h(X^{(i)}) \quad (5.17)$$

to the quantity $\mathbb{E}_\pi[h(X)]$, which is ensured by the Ergodic Theorem once we prove the ergodicity of the Metropolis chain.

Stationary distribution To see that π is the stationary distribution of the Metropolis chain, we first examine the Metropolis kernel. This kernel can be written as [84]

$$K(x, y) = \rho(x, y)q(y|x) + [1 - r(x)]\delta_x(y),$$

where $r(x) \equiv \int \rho(x, y)q(y|x)dy$ and δ_x denotes the Dirac mass in x . The first term in this kernel is the probability density of proposing a change of state from x to y and then accepting the proposed change. The second term represents the probability density of rejecting the candidate state, and therefore remaining in the current state.

Theorem 6. *Let $(X^{(i)})$ be a chain produced by Algorithm 1. For every conditional distribution q , the kernel of the chain satisfies the detailed balance condition with π , the target density.*

Proof. It is straightforward to verify that

$$\rho(x, y)q(y|x)\pi(x) = \rho(y, x)q(x|y)\pi(y)$$

and

$$[1 - r(x)]\delta_x(y)\pi(x) = [(1 - r(y)]\delta_y(x)\pi(y),$$

which together establish the detailed balance for the Metropolis-Hastings chain. \square

From Theorem 5, it immediately follows that π is also the stationary distribution.

Convergence properties We still need to assess the actual convergence properties of the Metropolis chain. We will show that the chain is ergodic, and, furthermore, that the empirical averages (5.17) converge to the expected value $\mathbb{E}_\pi(h)$.

The Metropolis-Hastings Markov chain is irreducible, since the conditional density q is positive:

$$q(y|x) > 0 \quad \forall (x, y) \in \mathcal{X} \times \mathcal{X}, \quad (5.18)$$

because it then follows that every subset of \mathcal{X} with positive Lebesgue measure can be reached in a single step. The M-H chain is also aperiodic, as it allows events such as $\{X^{(t+1)} = X^{(t)}\}$; that is, a proposed move of the chain is not always accepted, so that

$$P \left[\pi(X^{(t)})q(Y^{(t)}|X^{(t)}) \leq \pi(Y^{(t)})q(X^{(t)}|Y^{(t)}) \right] < 1. \quad (5.19)$$

Finally, since the density π is invariant for the chain, the chain is also positive, and Proposition 1 implies that it is recurrent.

Lemma 1. *If the Metropolis-Hastings chain $(X^{(i)})$ is irreducible under the probability measure π (i.e. it is π -irreducible), it is Harris recurrent.*

We shall omit the proof. We can summarize the above conclusions by stating that the M-H chain is ergodic, and we have the following theorem.

Theorem 7. *If the Metropolis-Hastings Markov chain $(X^{(i)})$ has a conditional density that satisfies (5.18) and (5.19), then*

(i) *the chain converges to the target distribution:*

$$\lim_{n \rightarrow \infty} \left\| \int K^n(x, \cdot) \mu(dx) - \pi \right\| = 0$$

(ii) *if $h \in L^1(\pi)$, then*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T h(X^{(i)}) = \int h(x) \pi(x) dx.$$

These are precisely the properties of the M-H chain that we are interested in, since they allow us to generate samples from π , and calculate its moments. Thus, instead of sampling directly from the target distribution, which might be very difficult, or, indeed, analytically impossible, the M-H algorithm samples from the proposal distribution, *and then introduces a correction*, such that the Markov chain's distribution eventually converges to the actual distribution of interest. In order to allow the Markov chain to converge to the target distribution, one has to run the chain for an initial period, called the *burn-in*. Once convergence has been established, the M-H algorithm primarily generates samples from the objective density π . Only the post-convergence samples are used for the Monte Carlo estimation of quantities. To monitor the convergence of the chain, one can study the acceptance probability or time traces of the quantities of interest, which should become stationary. We will demonstrate this in Chapter 6.

Random walk chain For a Metropolis-Hastings simulation to be efficient, on the one hand the acceptance probability throughout the simulation should be sufficiently high, but on the other hand the Markov chain still has to be able to move through the entire support of the target distribution π . A *random walk* chain is a common choice that satisfies these requirements. A Markov chain performs a random walk when candidate states $X^{(t+1)}$ are given by

$$X^{(t+1)} = X^{(t)} + \epsilon_t,$$

where ϵ_t is a (small) perturbation with distribution g , independent of $X^{(t)}$. The proposal distribution $q(y|x)$ is now of the form $g(y - x)$.

δ	0.1	0.5	1.0
Mean	0.339	-0.111	0.10
Variance	0.698	1.11	1.06

Table 5.1: Monte Carlo estimates of the mean and variance of a normal distribution $N(0, 1)$ based on a sample (after burn-in) from a Metropolis chain using a random walk on $[-\delta, \delta]$.

Common distributions for g include the uniform, the normal, Student's t and the Cauchy distributions.

As an example of a random walk, we consider the sampling of the normal distribution $\pi = N(0, 1)$ based on the uniform distribution on $[-\delta, \delta]$. The value of δ determines the largest possible Markov jumps. Since g is symmetric in this case, we can use the Metropolis algorithm, where the acceptance probability is given by $\rho(x, y) = \exp[(x^2 - y^2)/2]$. Figure 5.2 shows the result of three simulations of 15,000 samples each, produced by this method for $\delta = 0.1, 0.5$ and 1.0 respectively. The corresponding Monte Carlo estimates of the mean and variance are provided in Table 5.1. Figure 5.2 clearly shows the different speeds of convergence of the sampled averages as a function of δ . In this case, increasing δ leads to an improved correspondence of the histograms with the histogram of a normal distribution, due to a faster exploration of the support of π .

5.2.4.2 The multivariate Metropolis-Hastings algorithm

The univariate Metropolis-Hastings algorithm, as outlined above, can easily be extended to the multivariate case. In sampling a value for each random variable of interest, one can make use of the most recent already sampled values of the other variables. In the multivariate setting we will change our notation slightly, in order to conform to the notation used in Bayesian inference. As such, we will work with a data vector \mathbf{x} and a parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$, containing p parameters of interest. The samples that are generated are samples from the respective marginal distributions of the parameters.

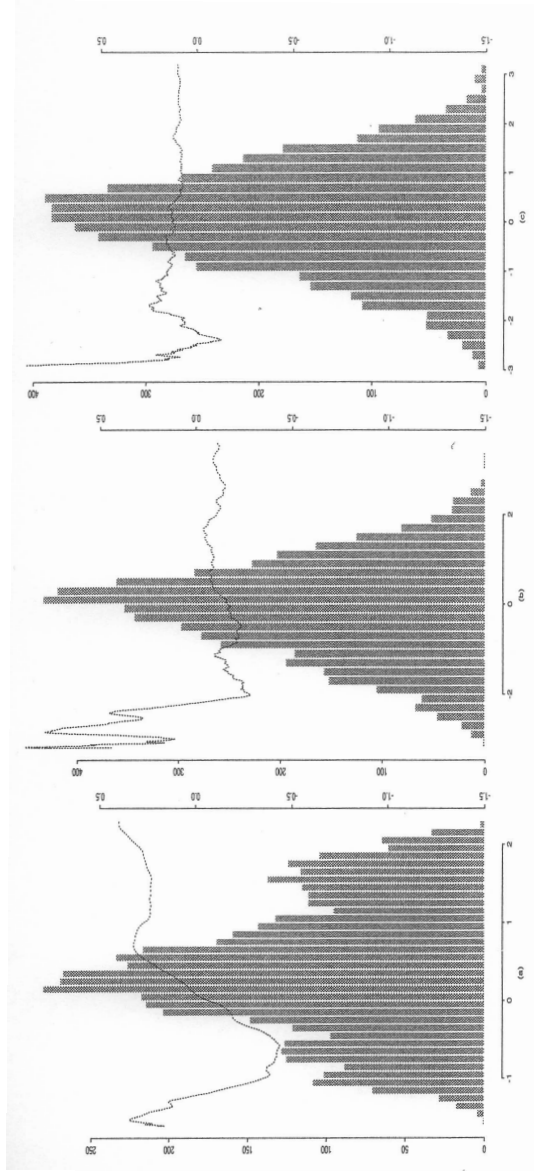


Figure 5.2: Histograms (after burn-in) for three random walk Metropolis simulations on $[-\delta, \delta]$ with (a) $\delta = 0.1$, (b) $\delta = 0.5$ and (c) $\delta = 1.0$. The Monte Carlo estimates of the means throughout the simulation have been superimposed (scale on the right side of the graphs).

5.2.4.3 The Gibbs sampler

While the Metropolis-Hastings algorithm is employed in the case where direct sampling from the target distribution is not feasible, the Gibbs sampler, also a Markov Chain Monte Carlo simulation method, can be used for direct sampling. We present the Gibbs sampler immediately in its multivariate form. It is required to have available the full univariate conditional distribution π_j for each parameter θ_j in an analytically closed form.

Algorithm 2 (Gibbs sampler).

```

1: Given  $\boldsymbol{\theta}^{(t)} = (\theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_p^{(t)})$ ,
2: repeat
3:    $\theta_1^{(t+1)} \sim \pi_1(\theta_1 | \theta_2^{(t)}, \dots, \theta_p^{(t)});$ 
4:    $\theta_2^{(t+1)} \sim \pi_2(\theta_2 | \theta_1^{(t+1)}, \theta_3^{(t)}, \dots, \theta_p^{(t)});$ 
    $\vdots$ 
5:    $\theta_p^{(t+1)} \sim \pi_p(\theta_p | \theta_1^{(t+1)}, \dots, \theta_{p-1}^{(t+1)}).$ 
6: until Satisfied

```

Note that in a single iteration of the Gibbs sampling algorithm, again always the most recent sample is used for every parameter. The Gibbs sampler does not involve an Accept/Reject step, as does the M-H algorithm.

The Gibbs sampler can be seen as a special case of the M-H algorithm, by considering it as a composition of p Markov chains.

Theorem 8. *The Gibbs sampling method of Algorithm 2 is equivalent to the composition of p Metropolis-Hastings chains, with acceptance probabilities uniformly equal to 1.*

We omit the proof.

Errors using inadequate data are much less than those using no data at all.

Charles Babbage

The purpose of models is not to fit the data but to sharpen the questions.

Samuel Karlin

Some years ago I had a conversation with a layman about flying saucers—because I am scientific I know all about flying saucers! I said ‘I don’t think there are flying saucers.’ So my antagonist said, ‘Is it impossible that there are flying saucers? Can you prove that it’s impossible?’ ‘No’, I said, ‘I can’t prove it’s impossible. It’s just very unlikely.’ At that he said, ‘You are very unscientific. If you can’t prove it impossible then how can you say that it’s unlikely?’ But that is the way that is scientific. It is scientific only to say what is more likely and what less likely, and not to be proving all the time the possible and impossible.

Richard Feynman

6

Integrated estimation of Z_{eff} from bremsstrahlung and CX spectroscopy

6.1 Introduction

Large amounts of plasma diagnostic data are generated on tokamaks. Plasma pulses on TEXTOR typically last around six seconds, but on other machines the discharges can last easily tens of seconds, up to even a few minutes. On ITER, pulses of several hundreds of seconds are foreseen. On the other hand, the physical phenomena of interest can vary on timescales up to microseconds, and many diagnostics are

required to be capable of a time resolution of at least a few tens of milliseconds. These enormous amounts of data need to be processed in an efficient way. The diagnostic data are used for the calculation of the physical quantities of interest, which are not only necessary for the interpretation of the physics behind the experiment, but also for the safe and flexible operation of the machine. This process of the derivation of quantities of interest from the raw measurements, is referred to as *data validation and analysis*. It can involve complex analysis techniques and consistency checks among diagnostics, both at the level of the raw and processed data. Since at both levels the data are affected by uncertainties of a diverse nature, a probabilistic analysis of all available information seems appropriate. This brings us in the realm of *Integrated Data Analysis* (IDA, or BIDA: Bayesian Integrated Data Analysis), an emerging field in fusion data processing, where data from a set of diagnostics are combined, possibly together with information on an underlying physical model, in order to extract the most probable desired physical quantities. It is even possible to combine data from multiple machines to aid physics studies. This chapter outlines the principles of IDA, and gives the results of an integrated analysis for the estimation of a local Z_{eff} value from both data of bremsstrahlung and CX spectroscopy. We work in a Bayesian context, making extensive use of the techniques introduced in the previous chapter. As such, we perform a so-called *Bayesian Integrated Data Analysis* (BIDA). This analysis was conducted entirely in the frame of the current work. As noted before, the data that we use to test our methods are all taken from the CXS system on JET, described in Section 4.2. Related initial results where TEXTOR data are used, are reported in Refs. [95], [96], [97] and [98].

As a start, a more elaborate motivation is given for the need for IDA methods.

6.2 Motivation

Plasma diagnosis is an extremely complex field involving the physics of the plasma itself, as well as the interoperation of numerous optical, mechanical and electronic components, and control and computational software. Errors in the eventual physical quantity of interest are thus introduced due to uncertainties at many different levels in the chain of derivation. These include uncertainties in the employed physical

model of the system plasma–diagnostic, in the measurement process, and in the calibration of the diagnostic system. Therefore, validation of data with regards to underlying physical models and complementary measurements is a crucial aspect of plasma diagnosis.

Although over the years there has been a steady development of increasingly sophisticated diagnostic techniques, validation of the data often remains a sideline activity, in particular if it comes to combining information from different sources. Nevertheless, as we shall demonstrate, the accuracy and consistency of measured data can be enhanced considerably, uncovering valuable information that was hidden in the data, provided enough effort is put in a systematic and rigorous data and error analysis.

In order to assess the impact of each of the uncertainties on the error bars of the quantities of interest, probability theory can be a useful tool. The idea is to model all uncertainties entering the analysis in terms of a suitable probability distribution, and from there construct a probability distribution of the desired physical quantity¹. In a Bayesian context, there is also the possibility to include expert prior knowledge. It is then possible to study the sensitivity of the physical quantity to the various uncertainties, and the most crucial uncertainties can be identified. Thus, all available information, of a heterogeneous nature, can be integrated into the analysis, including information on the physical model uncertainties, such as the mapping of plasma quantities onto magnetic surfaces derived from equilibrium calculations. Eventually, a ‘most probable’ estimate of the desired quantity can be calculated, and this will lead in many cases to more reliable (smaller error bars), more complete and more robust information about this quantity. An example is the derivation of the plasma electron density from interferometry on the one hand, and Thomson scattering on the other hand [99].

In a similar fashion, it is advantageous to integrate measurements from several diagnostics into a single probabilistic model. Indeed, if these measurements all contain information on the quantity of interest, the additional information from other diagnostic is likely to increase the accuracy of the desired physical quantities. Moreover, the results are less inclined to depend crucially on individual uncertainties related to one of the diagnostic methods.

The use of IDA can be motivated even more when the calculations

¹This is actually an extension of the common theory of Gaussian error propagation, applied by students in many physics practicals.

of a physical parameter from several diagnostics do not mutually correspond to satisfaction. An example is the subject of the current work: the inconsistency between Z_{eff} determined from bremsstrahlung measurements on the hand, and from a weighted summation of impurity concentrations on the other hand. In such a case, the question rises which one of the calculated estimates is actually the closest to the real underlying value. However, rather than weighing the results of the different diagnostics against each other, it is better to integrate all available information from the various diagnostics for the derivation of a single estimate *that is consistent with all data sets*. This requires the diagnosticians to:

1. identify all possible sources of uncertainty in the derivation of the desired quantities from the raw data,
2. propose estimates for the uncertainties.

Finally, the available space for diagnostic setups at ITER, and future fusion reactors, will be restricted, and physical quantities will need to be assessed from a limited data set. As such, any type of available information will have to be exploited. The concept of an Integrated Data Analysis provides an outstanding framework that can accomplish these tasks.

The possibilities of IDA that were mentioned in this paragraph will be demonstrated below for the estimation of Z_{eff} .

6.3 The BIDA recipe

In this section, we outline the basic recipe that can be used for the Bayesian integrated analysis of (tokamak) diagnostic data. An inherent feature of this approach is that it can be conducted through several stages, gradually increasing the complexity of the probabilistic model. As such, it is imperative to first identify the major sources of error in the quantities of interest, given the measured data, and rank these uncertainties in an order of increasing importance. Then, the uncertainties have to be formally quantified in terms of a suitable probability distribution. Statistical uncertainties in the data (measured data and calibration data) have to be modelled through likelihood PDFs, describing the error statistics of the measurement. Systematic uncertainties can be modelled by introducing nuisance parameters, possibly in a

hierarchical model using hyperparameters. The nuisance parameters subsequently have to be integrated out (marginalized). Uncertainties in the physical model itself can be taken into account by introducing and marginalizing proper model flexibility. Although also single experimental setups can be analyzed in this way, we here consider mainly the integration of multiple diagnostics.

First, estimates have to be determined for the statistical measurement errors. The effect of the statistical uncertainties is then modelled into an appropriate likelihood PDF. Often the normal distribution is selected here. In a first stage, when little is known about the exact nature of the various uncertainties, their contribution to measurement uncertainty can be gathered into only a few likelihood PDFs. Likewise, only a few general nuisance parameters can be used to describe the effect of all systematic uncertainties. In a later stage, the contribution of each individual uncertainty can be modelled in more detail. This is the approach taken in the present work, and it is a natural way of working, if one considers that virtually *any* uncertainty is made up of multiple ‘sub-sources’ of uncertainty.

The next step is to formalize any additional available information into a prior PDF. If one wants to make the least assumptions possible, one has to use uninformative priors.

Then all PDFs must be combined according to Bayes’ theorem. Bayes’ theorem specifies how to integrate all relevant information in a probabilistic framework in order to infer the quantities of interest. The nuisance parameters, if any, have to be marginalized, integrating out systematic effects and physical model uncertainties.

Finally, one ends up with the joint posterior PDF of the quantities of interest. One can then sample from the marginal distributions for each of the parameters, or calculate the marginal moments. In the present context, this (together with the marginalization of nuisance parameters) is mainly done using MCMC methods. Characteristics such as the mode or mean of the marginal distributions can then be used as best estimates for the respective parameters, and the variance can be used to construct statistical error bars.

The shape of posterior distributions allows to detect inconsistencies between different diagnostics. This can be signaled, for example, by a multimodal posterior distribution, or by a negligible overlap in parameter space of distributions obtained from using the data sets from the contributing diagnostics separately. It is usually the result of a system-

atic uncertainty that is not properly taken into account, or a too crude model simplification.

Next to the estimation of physical quantities, the posterior PDF allows also to study the impact of various experimental and model uncertainties on the error bars of these physical quantities. This process is named a *sensitivity analysis*. In practice, it involves reducing, or entirely switching off, individual uncertainties, and observing the effect on the posterior PDF. Thus, BIDA permits the optimization of experimental setups. Here, also the techniques of *Bayesian Diagnostic Design* can be of assistance, allowing even the optimization of future diagnostics. The idea is to maximize a utility function that quantifies the information gain on the quantities of interest as measurements are being taken. More information on this scheme can be found e.g. in Ref. [100].

In this work, only the Bayesian estimation of physical quantities (in particular Z_{eff}) is treated. Additional information on BIDA can be found under Refs. [101], [102], [103] and [99].

6.4 IDA for the estimation of Z_{eff} : local model

We will now propose a model that will allow us to make inferences about the plasma effective charge Z_{eff} , based on measurements of bremsstrahlung emissivity on the one hand, and data from CXS on the other hand. Many different techniques have been proposed and tested in the frame of the current work, before we arrived at this particular model and associated formalism. The idea behind the model stems from an analysis using a technique from statistical data processing known as *Independent Component Analysis* (ICA). This analysis was conducted also in the frame of the present work. Here, a relative time trace for the plasma-averaged Z_{eff} is recovered from a set of line-integrated bremsstrahlung emissivity measurements. ICA is a sort of *Blind Source Separation* (BSS) method, where a set of signals of interest (*sources*) is estimated from a set of measured *mixtures* [104]. In BSS, the details of the mixing process are also assumed to be unknown. An example is the so-called *cocktail-party problem*, where several people are speaking simultaneously in the same room, the combined speech signal being measured by a set of microphones, and the goal is to recover the individual speech signals. However, since the method does not allow to make inferences on the absolute value of the local Z_{eff} , we have not

included the details here, which can be found in Refs. [105], [106], [107] and [108].

The model that we propose here, initially includes both the *on-axis* Z_{eff} and the on-axis electron density n_e as unknown parameters to be estimated. Indeed, a typical electron density profile is also affected by a considerable amount of uncertainty, especially towards the plasma boundary. This is mainly due to the relatively strong arbitrariness that is involved in the inversion of only a few line-integrated measurements to a density profile, as well as the choice of a spline basis, which may very well lead to an overly smooth profile. Our measurements are not raw, unprocessed data, but rather two artificially constructed plasma quantities, ϵ and δ , defined below. Both ϵ and δ are the result of a long chain of calculations, involving many uncertainties. We will summarize these uncertainties in only a few variance and ‘nuisance’ parameters. This represents an initial approach towards the integrated estimation of Z_{eff} . The advantage is that our data descriptive (forward) model is simple. It is possible to increase the sophistication of the model, by considering for example the raw measurements of the respective detectors as input data. On the other hand, this will lead to a very complicated forward model. This is discussed in somewhat more detail in the general outlook in the concluding chapter, Chapter 7. We also want to note explicitly that we will start our analysis on measured and to be estimated quantities *on the magnetic axis* of the machine (JET). For the bremsstrahlung measurements, this implies that some sort of inversion procedure has been carried out. In Section 6.5, we will consider the estimation of Z_{eff} *profiles* from line-integrated bremsstrahlung measurements, and (local) CX measurements.

We start by viewing the model in the context of a BSS problem, and we consider the following ‘mixtures’ [105]:

- for the *local* bremsstrahlung emissivity ϵ_{ff} (see Equation (3.3)):

$$\epsilon = n_e^2 Z_{\text{eff}},$$

where

$$\epsilon \equiv \frac{\epsilon_{\text{ff}} \sqrt{T_e}}{C \bar{g}_{\text{ff}}}.$$

- for the *local* impurity densities n_i (see Equation (3.5)):

$$\delta = n_e (Z_{\text{eff}} - 1),$$

with

$$\delta \equiv \sum_i Z_i(Z_i - 1)n_i. \quad (6.1)$$

The sum is over all impurity species i .

Summarizing, our model is:

$$\begin{cases} \epsilon = n_e^2 Z_{\text{eff}} \\ \delta = n_e(Z_{\text{eff}} - 1) \end{cases}. \quad (6.2)$$

This constitutes a nonlinear source separation model, where ϵ and δ are the observed mixture signals, while n_e and Z_{eff} are the unknown sources. This brings us quite far from the original idea of BSS, as we should now view ϵ and δ as the signals picked up by the ‘microphones’, while n_e and Z_{eff} represent the ‘speakers’. Moreover, in this case the mixing process of the sources is clearly known, so there is no need to complicate the problem by using BSS methods.

Therefore, it is better to view the problem of determining Z_{eff} and n_e from the perspective of Bayesian estimation, where we can keep working with ϵ and δ as the two measured data, and where Z_{eff} and n_e are the parameters of interest. Indeed, as already stated, initially we will treat n_e also as an unknown. However, the quantity δ has a dependency on n_e through the beam attenuation, and the LIDAR density is used here as a measurement². Therefore, at this stage, there is no real practical use in estimating n_e from ϵ and δ as well. On the other hand, treating n_e as an unknown will allow us to illustrate many concepts that will be of prime importance in the Bayesian scheme for the estimation of a consistent Z_{eff} . In addition, we can compare the estimated n_e to the density measured by other diagnostics, for example through LIDAR. The on-axis n_e measurement has a small error bar (Section 4.3). Therefore, we will for the time being take the degree of correspondence of our estimated density to the LIDAR density, as a measure to judge the validity of our Z_{eff} estimate as well. This should be seen also as a function of the model complexity: we might be prepared to put more trust in the estimated Z_{eff} values as the working model sophistication

²On JET, the density profile mapped on flux surfaces is derived from LIDAR Thomson scattering measurements (see Ref. [29]), which are normalized against interferometric measurements. In short we will refer to the density measurements as the ‘LIDAR density’.

Discharge #	Programme
59193, 59194, 59186	Long distance high power Lower Hybrid Current Drive coupling on plasmas with Internal Transport Barrier
60718	High confinement, high density ELMy H-modes (+ feedback controlled impurity seeding)
61352	Study of D and T fueling for ITB plasmas

Table 6.1: Scientific programme associated with the JET discharges under study in this work.

is increased. In any case, an estimated n_e that corresponds well to the LIDAR n_e , is an indication that at least our calculations and coding are sound (the reverse is not necessarily true). At a later stage, we will also explicitly incorporate electron density measurements from LIDAR into the model, where we will cover at least the statistical uncertainties on the density measurements. Finally, the models described in this work can be extended to estimate not only Z_{eff} , but also n_e and other plasma quantities from a set of *raw* measurements. As such, it makes sense to already now demonstrate the possibilities of estimating multiple plasma quantities of interest from a limited set of measurements. For example, the off-axis electron density can be affected by many uncertainties too, especially when it comes to the reconstruction of a density profile. Therefore, also the determination of n_e can benefit from an integrated analysis.

Incidentally, any uncertainty in the LIDAR density will propagate into the uncertainty on δ through its n_e -dependence. Thus, part of the uncertainty on δ is made up of uncertainty on the LIDAR measurements.

We will continue to work with the model (6.2), modelling several statistical (Section 6.4.2) and systematic (Section 6.4.3) uncertainties in due time. Since this model involves only local, on-axis quantities, we will refer to it as the *local model*.

Finally, Table 6.1 gives for all JET discharges used in this work some information on the associated scientific programme.

6.4.1 Inversion of the local model

The model (6.2) is a system of two equations, with two data (ϵ and δ) and two unknowns (n_e and Z_{eff}). If we assume that no measurement error was made, we can solve this system exactly. Hence, we presume also a zero error on n_e and Z_{eff} . The solution can already learn us a lot about the more realistic problem, where measurement error is present.

After some simple algebra, we find the following solutions of (6.2):

$$Z_{\text{eff,inv}} = \frac{2\epsilon + \delta^2 + \delta\sqrt{4\epsilon + \delta^2}}{2\epsilon} \quad \vee \quad Z_{\text{eff,inv}} = \frac{2\epsilon + \delta^2 - \delta\sqrt{4\epsilon + \delta^2}}{2\epsilon} \quad (6.3)$$

$$n_{e,\text{inv}} = \sqrt{\frac{\epsilon}{Z_{\text{eff,inv}}}} \quad (6.4)$$

It can be seen quite easily that the second solution for Z_{eff} in (6.3) is always smaller than 1, so we will no longer consider it, nor its associated solution for n_e .

We will calculate the solutions (6.3) and (6.4) from measurements of ϵ and δ for two JET discharges where only fully stripped carbon was monitored by the CX system (see also Section 3.2.5). It is only from Section 6.4.3 on, where also systematic uncertainties on the measurements are considered, that the influence from other impurities can be taken into account under some circumstances. In order to be able to visualize the results better, we work with time traces of measurements ϵ and δ , and we also estimate n_e and Z_{eff} time traces. The number of input data is then still two, since we will process all time points in parallel.

We have selected the JET discharges, #60718 and #61352, for which the time evolution of the on-axis continuum and CX Z_{eff} was already shown in Figures 3.5 and 3.6, respectively. Figure 6.1 shows the evolution in time for discharge #60718 of the on-axis n_e , T_e , ϵ_{ff} and n_C , during operation of the neutral beam in JET octant 8. The signals for the derived parameters ϵ and δ are also shown. Figure 6.2 gives a similar overview for pulse #61352. The error bars shown correspond to the relative errors given in Table 4.2. The error bars on ϵ and δ have been calculated assuming that the errors on the other quantities are of a pure statistical nature, signifying a single standard deviation. We will have to come back to this assumption later on. This yields 20% for the relative error on ϵ and 35 % for δ . Figure 6.3 now shows for pulse

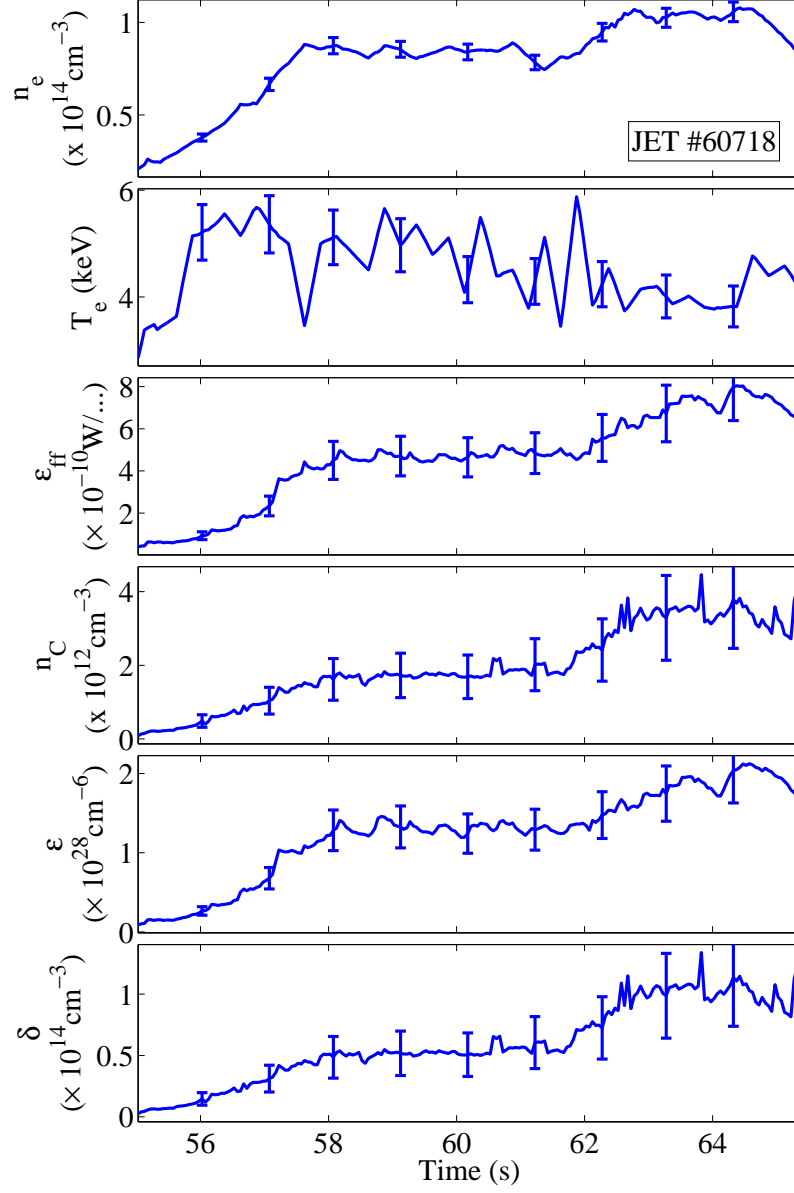


Figure 6.1: Time traces for the on-axis n_e , respectively T_e , ϵ_{ff} , n_C , ϵ and δ for JET pulse #60718. The units for ϵ_{ff} have been abbreviated, and are $\text{W} / (\text{cm}^2 \text{ sr } \text{\AA})$.

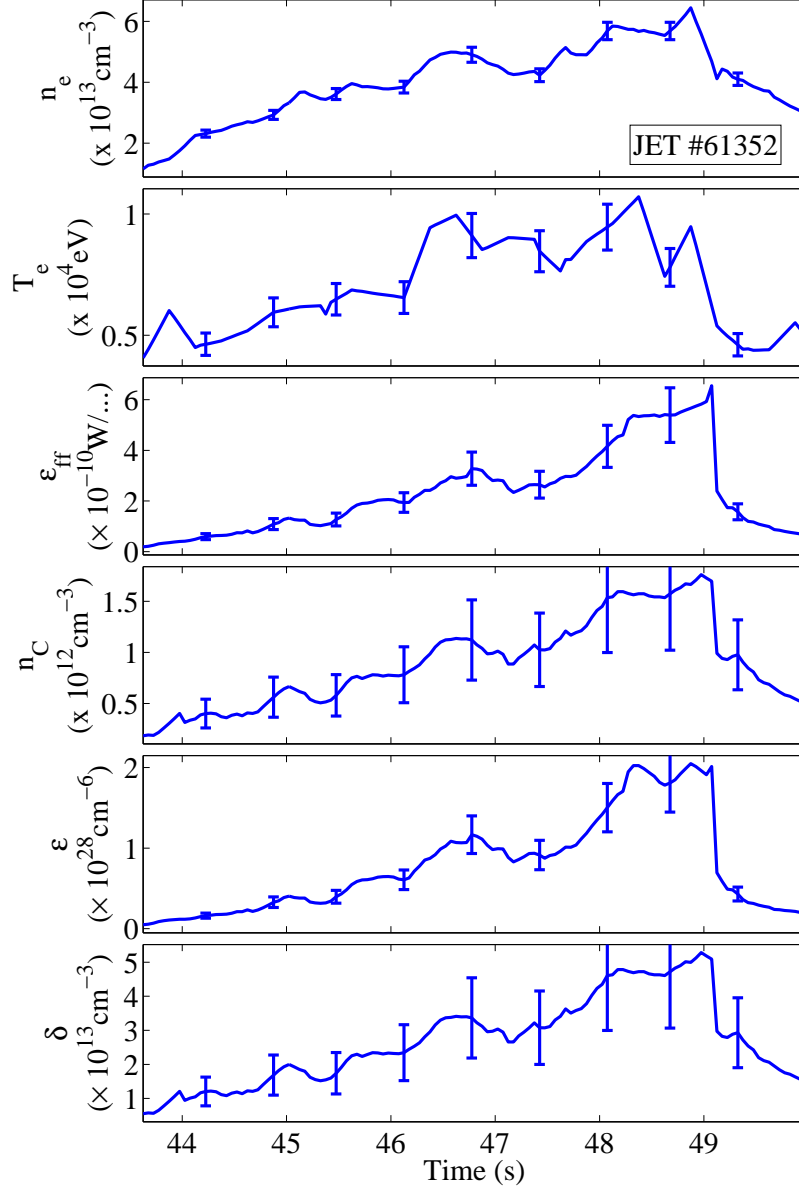


Figure 6.2: Time traces for the on-axis n_e , respectively T_e , ϵ_{ff} , n_C , ϵ and δ for JET pulse #61352. The units for ϵ_{ff} have been abbreviated, and are $\text{W} / (\text{cm}^2 \text{ sr } \text{\AA})$.

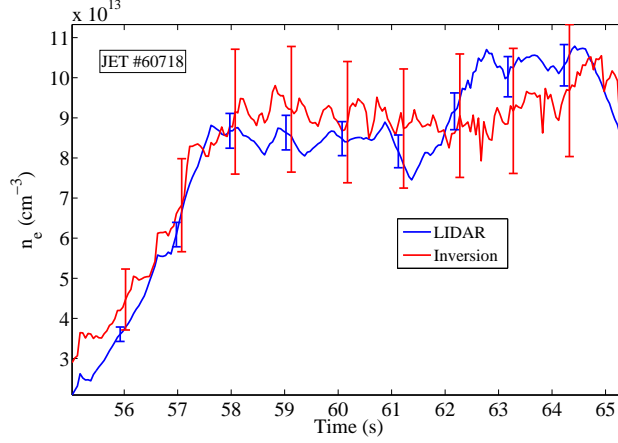


Figure 6.3: Time traces for the on-axis density from LIDAR and the calculated density through model inversion for pulse #60718.

#60718 the solution $n_{e,\text{inv}}$ obtained from (6.4), and the comparison with the signal for n_e measured by LIDAR. The signal for $n_{e,\text{inv}}$ is for the most part similar to the density signal from LIDAR.

Figure 6.4 shows the continuum and CX Z_{eff} signals, and the solution $Z_{\text{eff,inv}}$. It is rather striking that the solution is much more similar to the CX Z_{eff} than to the continuum Z_{eff} . We will go further into this behaviour later on.

Both figures also show the error bars for the signals, the width of which we define again as a single standard deviation. The error bars were calculated by simple Gaussian error propagation analysis from the errors on ϵ and δ . The relative error for the solution $n_{e,\text{inv}}$ is about 17%: considerably larger than for the density from LIDAR (5%). The error bars for the two n_e estimates overlap two-by-two. The relative error for the calculated Z_{eff} from model inversion is about 25%, the same as for the continuum Z_{eff} , and slightly higher than the error bar on the CX Z_{eff} . Again the error bars for the three Z_{eff} estimates overlap two-by-two. Therefore, although the continuum Z_{eff} and the CX Z_{eff} signals are quite different, particularly in qualitative behavior, the error analysis shows that these two Z_{eff} estimates are still mutually consistent within the defined error bounds. Still, there could be a case where the underlying physical mechanisms governing the value of Z_{eff} suggest a moderate variation of Z_{eff} (e.g. like discussed in Section 4.1.4.3). It

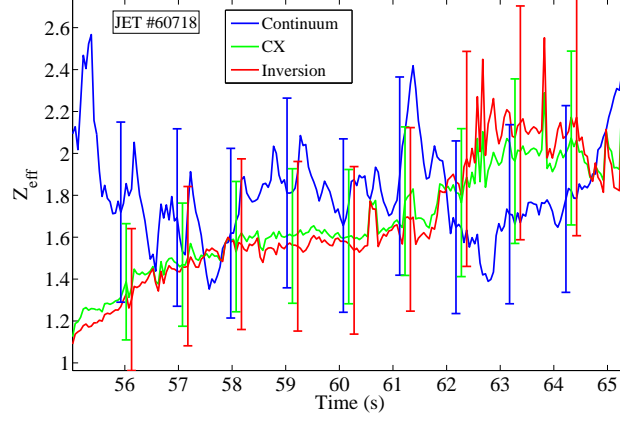


Figure 6.4: Time traces for the on-axis continuum Z_{eff} and CX Z_{eff} , and the calculated Z_{eff} through model inversion for pulse #60718.

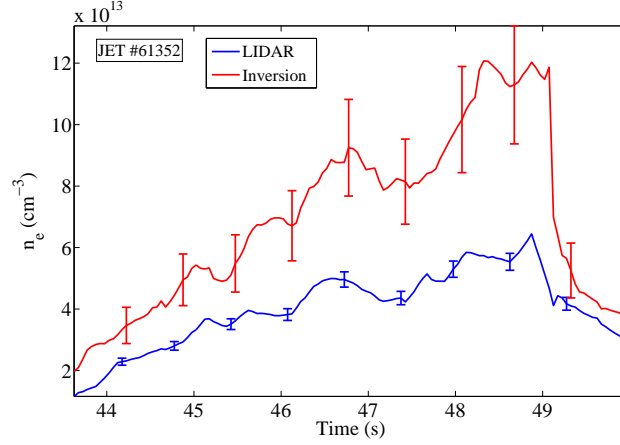


Figure 6.5: Time traces for the on-axis density from LIDAR and the calculated density through model inversion for pulse #61352.

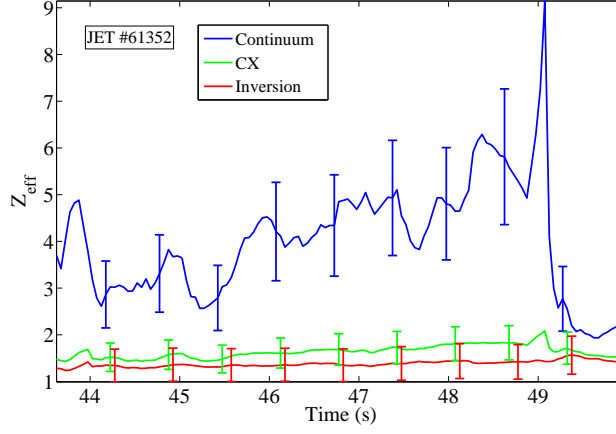


Figure 6.6: Time traces for the on-axis continuum Z_{eff} and CX Z_{eff} , and the calculated Z_{eff} through model inversion for pulse #61352.

would then be very well possible that the results for the continuum and the CX Z_{eff} are different to the extent that we are unable to support this physical hypothesis with the experimental data. Indeed, a priori we attach more or less the same value to the continuum and CX Z_{eff} estimates. In this sense, we might still feel that we need a higher degree of confidence in the Z_{eff} values, or, put differently, that the continuum and CX Z_{eff} estimates should correspond to a higher degree.

The situation is different for JET pulse #61352, where the systematic difference between the continuum and CX Z_{eff} is much larger. Observing Figures 6.5 and 6.6, it can be noticed that the solution $n_{e,\text{inv}}$ for the density now matches much less the density from LIDAR, whereas the solution for Z_{eff} is still very close to the CX Z_{eff} . However, in this case the error bars of the continuum and CX Z_{eff} do not overlap two-by-two. This corresponds to our intuition, showing that for this particular discharge the continuum and CX Z_{eff} are clearly totally inconsistent, exhibiting strong systematic discrepancies. This will be elaborated on in a while.

The direct inversion of the system of equations (6.2) is an instructive exercise, and the n_e and Z_{eff} values that are thus calculated are the first to be consistent with both our bremsstrahlung and CX data sets. However, we have no measure that indicates that these estimates are any better than what we had before. In other words, the inversion

does not bring us much closer to our aim of producing a more reliable, consistent Z_{eff} .

6.4.2 Inference from the local model: statistical uncertainties

6.4.2.1 Model definition

As a next step, in order to make the model somewhat more realistic, we will take into account some uncertainties on the data. We have again taken two measurements ϵ and δ , and we wish to estimate from this the n_e and the Z_{eff} that are consistent with both measurements. To begin with, we will again assume that all uncertainties in the data are of a statistical nature, not systematic. We will model all statistical uncertainties entering the data through two uncertainty terms ν_1 and ν_2 :

$$\begin{cases} \epsilon = n_e^2 Z_{\text{eff}} + \nu_1 \\ \delta = n_e(Z_{\text{eff}} - 1) + \nu_2 \end{cases}, \quad (6.5)$$

where $\nu = [\nu_1, \nu_2]^T$, which we assume to be distributed independently identically Gaussian:

$$\nu \sim \mathcal{N}(0, \Sigma_\nu), \quad \Sigma_\nu \equiv \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix},$$

where γ_i denotes a precision. The motivation for a Gaussian noise distribution is the Central Limit Theorem. Indeed, the noise on a measurement is usually due to a large amount of small, independent influences. The sum of these random variables then approaches a Gaussian distribution. The fact that we take the noise on ϵ to be independent of the noise on δ , means that the measurements should have been performed with two different instruments. In our case, this is not really true, as we have used bremsstrahlung measurements from the background of the CX spectrum. However, since both ϵ and δ are not raw data, we expect the error statistics to be quite different. Therefore, we will assume no noise correlation as yet, an assumption that can be relaxed in the future.

In order to quantify the precisions γ_1 and γ_2 , we have used again a standard deviation of 20% on ϵ , and 35% on δ .

We will now construct the forward model in terms of appropriate likelihood functions. Since the noise terms ν_1 and ν_2 both have a normal

distribution and are independent, we can write the likelihood for the two unknown (and desired) parameters Z_{eff} and n_e as:

$$L(n_e, Z_{\text{eff}}) = p(\epsilon, \delta | n_e, Z_{\text{eff}}, I) = \left(\frac{\gamma_1}{2\pi}\right)^{1/2} \exp\left[-\frac{\gamma_1}{2}(\epsilon - n_e^2 Z_{\text{eff}})^2\right] \\ \times \left(\frac{\gamma_2}{2\pi}\right)^{1/2} \exp\left(-\frac{\gamma_2}{2}[\delta - n_e(Z_{\text{eff}} - 1)]^2\right), \quad (6.6)$$

where I represents any additional information we might possess. We will call the first likelihood (containing ϵ) the *continuum likelihood*, and the second likelihood (containing δ) the *CX likelihood*. If we have a series of measurements ϵ_i and simultaneous δ_i at a series of time points, we can incorporate them into a single likelihood function:

$$L(\epsilon, \delta | n_e, Z_{\text{eff}}, I) = \left(\frac{\gamma_1}{2\pi}\right)^{T/2} \exp\left[-\frac{\gamma_1}{2} \sum_i (\epsilon_i - n_{e_i}^2 Z_{\text{eff}_i})^2\right] \\ \times \left(\frac{\gamma_2}{2\pi}\right)^{T/2} \exp\left[-\frac{\gamma_2}{2} \sum_j [\delta_j - n_{e_j}(Z_{\text{eff}_j} - 1)]^2\right], \quad (6.7)$$

where T is the total number of time points and we have introduced the shorthand notation \mathbf{s} for any time series s_1, \dots, s_T . We have also assumed that the noise on the measurements at different time points is uncorrelated; that is, at this point we do not model any time structure (which will prove to be not necessary anyway). Therefore, we can just as well work with a set of T likelihoods (6.6), as with the full joint probability density (6.7).

We do not wish to assume a lot of prior information on Z_{eff} and n_e , so we will use a relatively uninformative prior. Remembering the discussion in Section 5.1.5.1, we choose a uniform prior, cut off at appropriate boundaries. For TEXTOR, we take the following safe intervals:

$$n_e \in [0, 6 \times 10^{13} \text{cm}^{-3}], \\ Z_{\text{eff}} \in [1, 5]. \quad (6.8)$$

On JET, n_e can reach somewhat higher values, so we choose:

$$n_e \in [0, 15 \times 10^{13} \text{cm}^{-3}], \\ Z_{\text{eff}} \in [1, 5]. \quad (6.9)$$

Every value of Z_{eff} or n_e falling outside these respective intervals, is assumed to have zero probability of being realized.

Following Bayes' theorem, Equation 5.5, we can now write down a posterior distribution (for each time point), which matters only up to a proportionality factor:

$$\begin{aligned} p(n_e, Z_{\text{eff}} | \epsilon, \delta, I) &\sim p(\epsilon, \delta | n_e, Z_{\text{eff}}, I) p(n_e, Z_{\text{eff}}, I) \\ &\sim \exp \left[-\frac{\gamma_1}{2} (\epsilon - n_e^2 Z_{\text{eff}})^2 - \frac{\gamma_2}{2} (\delta - n_e [Z_{\text{eff}} - 1])^2 \right] \\ &\times U_{n_e} U_{Z_{\text{eff}}}, \end{aligned} \quad (6.10)$$

where U_{n_e} and $U_{Z_{\text{eff}}}$ are the uniform densities defined on the respective intervals (6.8) or (6.9).

We will also treat the case where the electron density measurements from LIDAR are considered to be given explicitly (not only through δ), which gives us three measurements. We will then still consider n_e as an unknown parameter, but for the problem of the determination of Z_{eff} , n_e then has become a nuisance parameter, which we have to integrate out of the posterior. In practice, the posterior changes through the addition of an extra likelihood factor:

$$\begin{aligned} p(n_e, Z_{\text{eff}} | \epsilon, \delta, n_{e,I}, I) &\sim \exp \left[-\frac{\gamma_1}{2} (\epsilon - n_e^2 Z_{\text{eff}})^2 - \frac{\gamma_2}{2} (\delta - n_e [Z_{\text{eff}} - 1])^2 \right. \\ &\quad \left. - \frac{\gamma_1}{2} (n_{e,I} - n_e)^2 \right] U_{n_e} U_{Z_{\text{eff}}}, \end{aligned} \quad (6.11)$$

with $n_{e,I}$ denote the LIDAR measurement, and γ_1 the associated precision. We will call the extra likelihood factor the *LIDAR likelihood*.

6.4.2.2 Posterior calculations

Measurements of ϵ and δ Suppose we have done a (series of) measurement(s) of ϵ and δ . Since we essentially search for the two unknown parameters n_e and Z_{eff} , we work in a two-dimensional parameter space. For problems of such low dimensionality, it is still computationally feasible to calculate the posterior density (6.10) on a (sufficiently large) grid of values for n_e and Z_{eff} . Figure 6.7 gives a contour plot of the posterior density resulting from such a procedure. Here, an ϵ and δ value were taken from pulse #61352 at 47s. The three-dimensional plot of the posterior is shown in Figure 6.8. In order to make inferences about n_e or Z_{eff} individually, we have to marginalize the joint posterior. In this two-dimensional problem, this integration can still be done with traditional computational methods. In this case, we have just summed over the parameter values on the (sufficiently fine) grid; the results are

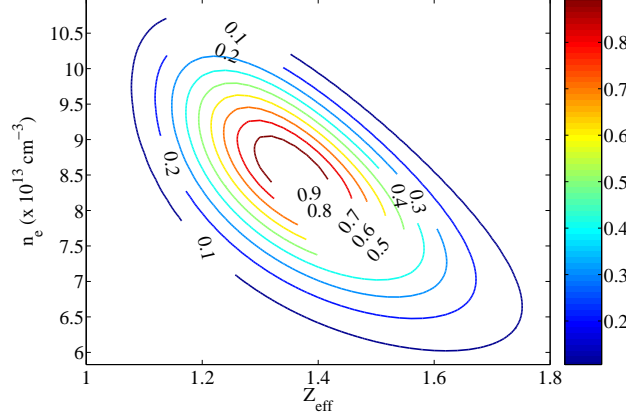


Figure 6.7: Contour plot for the joint posterior density of the on-axis n_e and Z_{eff} for JET #61352 at 47s. The numbers indicate probability density.

displayed in Figure 6.9. The maximum (i.e. the mode) of the marginal posteriors equal the respective solutions from the deterministic model (6.2). The maximum of the posterior distribution is referred to as the *Maximum a Posteriori* (MAP) estimate for the parameters of interest. A measure for the uncertainty at the maximum of the joint posterior is given by the covariance matrix. This can be approximated by taking the inverse Hessian at the maximum. If p denotes the posterior, the Hessian is given by

$$\begin{bmatrix} \frac{\partial^2 p}{\partial Z_{\text{eff}}^2} & \frac{\partial^2 p}{\partial Z_{\text{eff}} \partial n_e} \\ \frac{\partial^2 p}{\partial n_e \partial Z_{\text{eff}}} & \frac{\partial^2 p}{\partial n_e^2} \end{bmatrix}.$$

In effect, this comes down to approximating the joint posterior at the maximum by a Gaussian. The diagonal elements of the covariance matrix give the posterior variances on Z_{eff} , respectively n_e . Alternatively, one can calculate the standard deviation of the marginals. On the other hand, the form of the posterior reveals the correlation between Z_{eff} and n_e ³. The off-diagonal elements of the covariance matrix give the cross-correlation coefficients between Z_{eff} and n_e . However, the joint and marginal distributions give even much more information on the parameters n_e and Z_{eff} . Indeed, these probability distributions

³A multimachine database has yielded the following scaling relation between Z_{eff} and the line-averaged density \bar{n}_e : $Z_{\text{eff}} - 1 \sim 1/\bar{n}_e^2$ [109, 110].

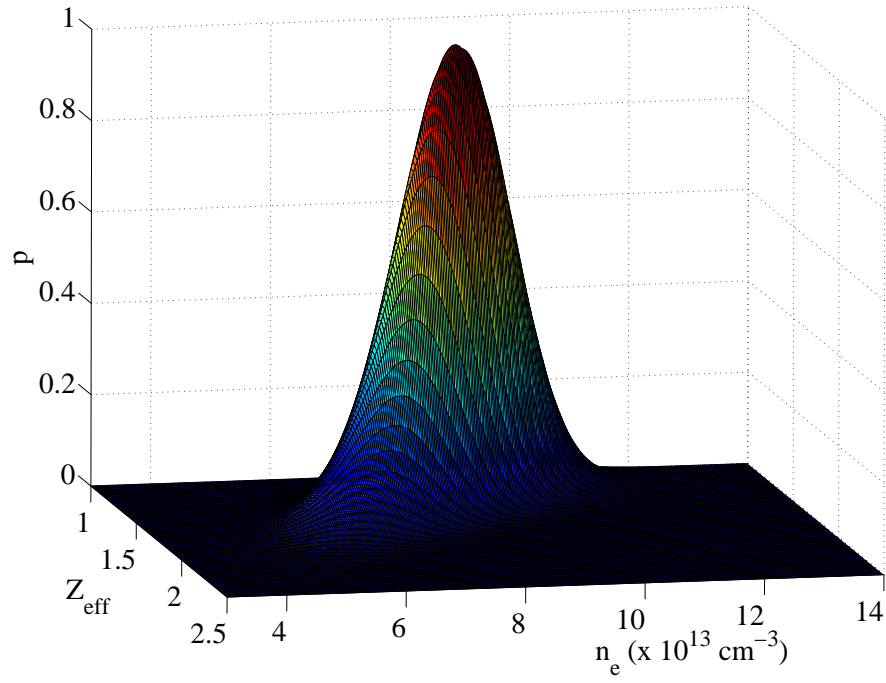


Figure 6.8: The joint posterior density of the on-axis n_e and Z_{eff} for JET #61352 at 47s. The numbers indicate probability density.

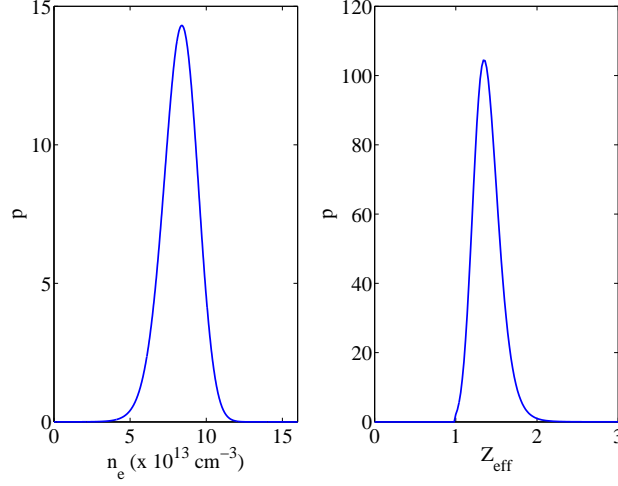


Figure 6.9: Marginal posterior distributions for the on-axis n_e and Z_{eff} for JET #61352 at 47s. Note how the probability density is cut off at $Z_{\text{eff}} = 1$.

yield the most complete picture one can obtain on the parameters of interest, given the model, the data and the error bars on the data. The probability distributions allow not only to calculate any moment of the distribution, but also any quantile. In addition, the joint posterior is quite skewed, and depending on the skewness, an estimate for the marginal means can differ substantially from an estimate of the marginal modes for n_e and Z_{eff} . We have calculated the marginal posterior mean and standard deviation for n_e and Z_{eff} at each time point. The results, shown in Figures 6.10 and 6.11, are not much different from the signals obtained via deterministic inversion of the model (i.e. the MAP estimates), although the respective error bars are somewhat smaller in the present case.

Measurements of ϵ , δ and $n_{e,I}$ We now turn to the case where not only ϵ and δ , but also an electron density measurement from LIDAR $n_{e,I}$ is given, each with its respective error bar. We perform the same analysis as in the previous subsection, again starting from the measurements of JET pulse #61352 at 47s. However, we now work with the posterior given in (6.11). We have immediately displayed the marginals for n_e and Z_{eff} in Figure 6.12. According to our expectations, and comparing

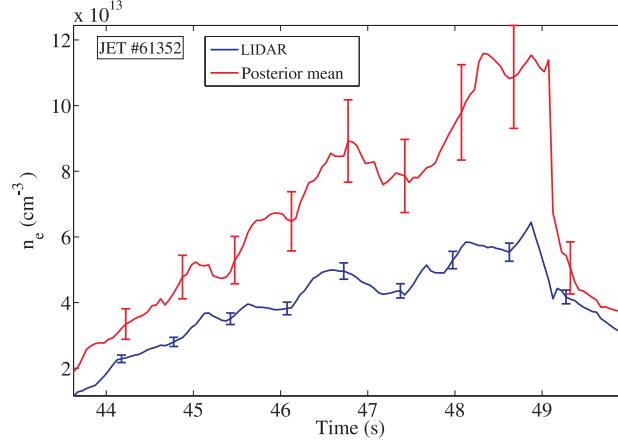


Figure 6.10: Marginal posterior means of the on-axis n_e for JET #61352. The error bars on the posterior means correspond to a single standard deviation.

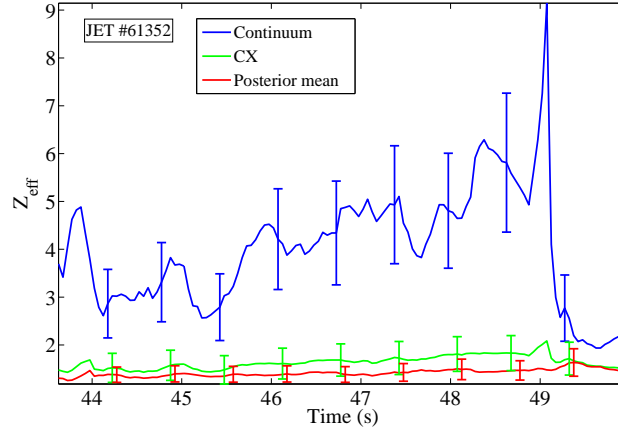


Figure 6.11: Marginal posterior means of the on-axis Z_{eff} for JET #61352. The error bars on the posterior means correspond to a single standard deviation.

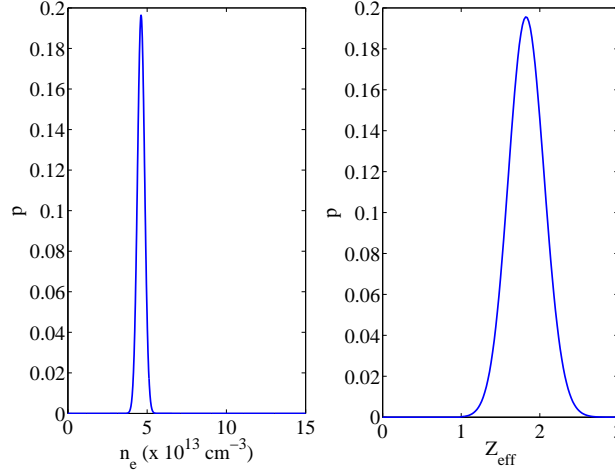


Figure 6.12: Marginal posterior distributions for the on-axis n_e and Z_{eff} for JET #61352 at 47s. The density from LIDAR is taken as an additional measurement.

to Figure 6.9, the characteristic width of the marginal posterior for n_e has shrunk considerably with the inclusion of the extra density measurement. The number of measurements has thus gone from two to three, and, in the line of the discussion in Section 5.1.6, this decreases the relative influence of the uniform prior on the posterior. Again, we have calculated the marginal posterior mean and standard deviation for n_e and Z_{eff} at each time point. The results are displayed in Figures 6.13 and 6.14. The estimate for n_e now matches the density from LIDAR much better than in the case where only the measurements ϵ and δ were used, with comparable error bars. Of course, this is largely due to the high trust we have chosen to put in the LIDAR (5% error bars). We will go further into this in the discussion below. To repeat the analysis at other radial positions, a higher uncertainty on $n_{e,I}$ should be taken into account. The estimate for Z_{eff} has risen somewhat to a level in between the continuum Z_{eff} and the CX Z_{eff} . It is however still much more similar to the CX Z_{eff} , than to the continuum Z_{eff} , both in time evolution and absolute value. In addition, one can note that the error bars on the estimated Z_{eff} are slightly smaller than for the continuum and CX Z_{eff} .

We have performed the analysis, including the $n_{e,I}$ measurements,

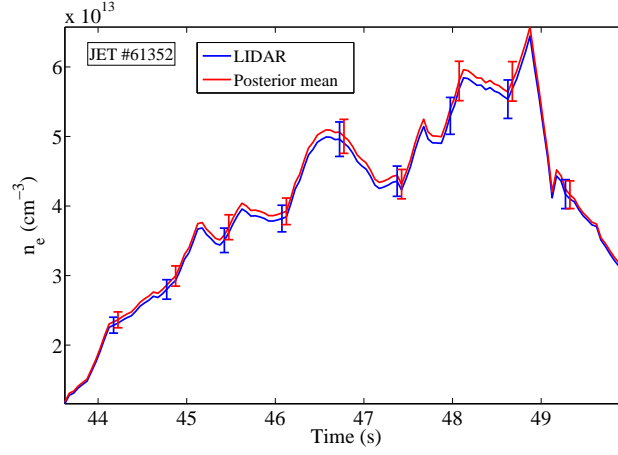


Figure 6.13: Marginal posterior means of the on-axis n_e for JET #61352, including also LIDAR measurements.

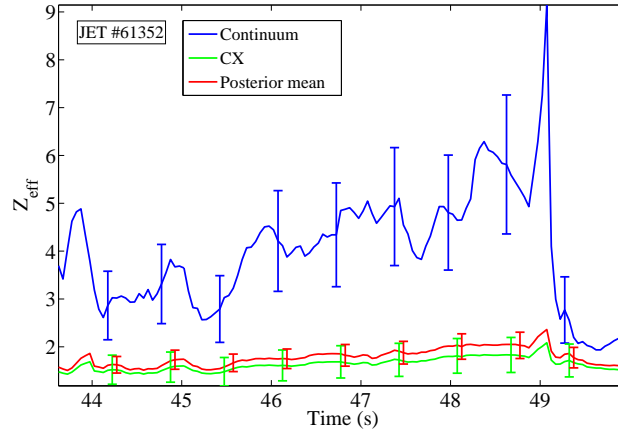


Figure 6.14: Marginal posterior means of the on-axis Z_{eff} for JET #61352, including also LIDAR measurements.

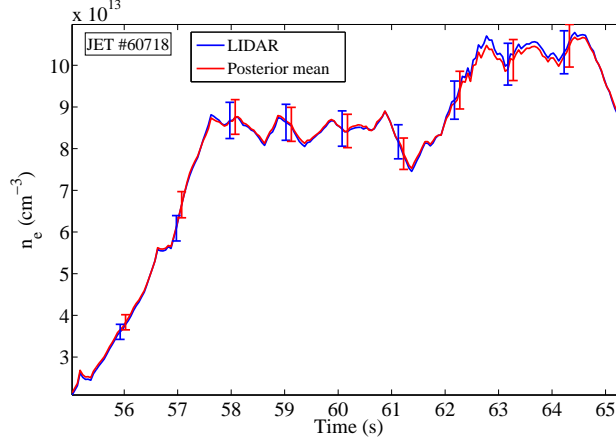


Figure 6.15: Marginal posterior means of the on-axis n_e for JET #60718, including also LIDAR measurements.

also for the JET discharge #60718. The estimates are shown in Figures 6.15 and 6.16. Again, the estimate for n_e corresponds very well with the LIDAR measurements. Notice also how the estimated Z_{eff} signal is similar to the CX Z_{eff} , but nevertheless bears some resemblance with the continuum Z_{eff} as well. The error bars of the estimated Z_{eff} are also substantially smaller than for the continuum and CX Z_{eff} .

Discussion First, we want to stress that for both discharges under study, the estimated Z_{eff} is more similar to the CX Z_{eff} *in spite of* the fact that the relative error on δ (35%) is higher than the one on ϵ (20%).

Next, as already mentioned, in discharge #61352 the continuum and CX Z_{eff} are very inconsistent. In this case, the estimated Z_{eff} is much more consistent with the CX Z_{eff} than with the continuum Z_{eff} . However, in discharge #60718, showing some better consistency between the continuum and CX Z_{eff} , the estimated Z_{eff} is also consistent with both, and when incorporating explicit LIDAR measurements, the estimate seems to take over at certain occasions within the discharge some of the behavior of the continuum Z_{eff} time trace. This again agrees with a, be it somewhat primitive, first idea that one might have about a Z_{eff} that is consistent with all data, namely that it should display some features of both the continuum and CX Z_{eff} .

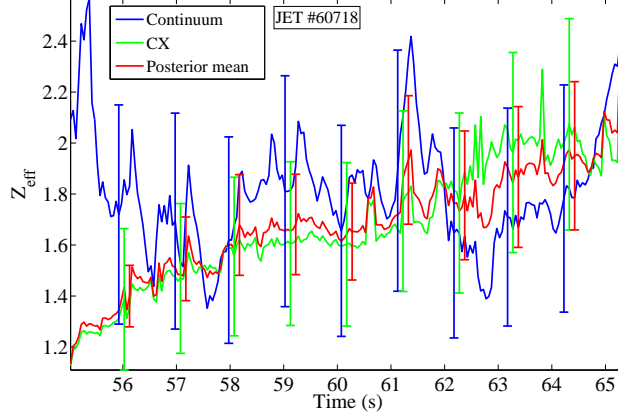


Figure 6.16: Marginal posterior means of the on-axis Z_{eff} for JET #60718, including also LIDAR measurements.

The continuum Z_{eff} can display sudden changes on relatively short time scales. Observe for example the rise and subsequent fall with a factor of about 1.5 between 61 s and 62 s in discharge #60718. One can verify that this peak is due to a simultaneous drop in n_e . The CX Z_{eff} also displays a peak at this point, but much less pronounced. The height of the corresponding peak in the estimated Z_{eff} is somewhere in between the height for the CX and for the continuum Z_{eff} . A similar example can be found at about 62.8 s, where, as a result of a peak in the density signal, the continuum Z_{eff} displays a sudden drop, which is followed to some extent by the estimated Z_{eff} . Thus, the density signal determines for a part the behavior of the estimated Z_{eff} , which is particularly striking when this behavior is not supported by the CX data. This is quite natural, since we were prepared to put a lot of trust in the LIDAR density values in the first place. Let us see what happens when we increase the uncertainty on $n_{e,I}$ from 5% to, say, 20%. Note that with such a high uncertainty on the density, it becomes quite infeasible to still calculate a meaningful Z_{eff} value from bremsstrahlung measurements alone. Indeed, the error on the continuum Z_{eff} goes up to 50% while the error on the CX Z_{eff} goes to 25%. Thus, one has to be relatively certain of the n_e measurements in order to calculate a sensible continuum Z_{eff} . An increased uncertainty on $n_{e,I}$ implicitly increases also the error on δ . At first, however, we have taken the error on δ to remain at 35%, so that a clear difference can be seen with our previous

analysis. The result for the Z_{eff} estimate in pulse #61352 is shown in Figure 6.17. It is notable that the estimated Z_{eff} now coincides almost perfectly with the CX Z_{eff} . This is a good sign, because when the error on n_e increases to 20%, we have no reason anymore to trust the continuum Z_{eff} at all. In addition, the error bars indicate that even the consistency of the estimated Z_{eff} with the continuum Z_{eff} has increased substantially.

If we now also increase the error on δ to, say, 50%, as a result of its n_e -dependence, we get the result displayed in Figure 6.18, showing again a good correspondence to the CX Z_{eff} , despite the large error on the δ measurement. Finally, we performed the same analysis, with an error of 20% on n_e and 50% on δ , for discharge #60718, shown in Figure 6.19. Again, any similarity of the estimated Z_{eff} to the continuum Z_{eff} has almost disappeared.

Thus, the more trust we are willing to put in the measured electron density, the more the estimated Z_{eff} resembles the continuum Z_{eff} . This is a consequence of the quadratic dependence of the continuum Z_{eff} on n_e , compared to the approximate linear dependence of the CX Z_{eff} on n_e . If we want the estimated Z_{eff} to display a good deal of the time behavior of the continuum Z_{eff} , we have to increase the certainty on the measured electron density almost to unrealistically high levels. However, since diagnosticians generally put a good trust in the on-axis electron density measurements, we will, in the following, keep an error of 5% on the $n_{e,I}$. The error on δ is then again the previous 35%.

Finally, the analysis in this section also makes clear that putting a high trust in a certain measurement, can influence the results of the probabilistic estimation. However, in general the exact values of the relative statistical errors on measurements, play a minor role in the models used here.

6.4.2.3 Consistency analysis

We would now like to visualize the degree of consistency of the various measured data sets. To this end, we have made an overview of the marginal posterior distributions for n_e and Z_{eff} , taking into account various (combinations of) data sets. As an example, we have taken again JET pulse #61352 at 47 s, because here the traditional continuum and CX Z_{eff} are very different. Figure 6.20 presents the marginal distributions for n_e for a posterior containing, respectively, a pure LIDAR likelihood,

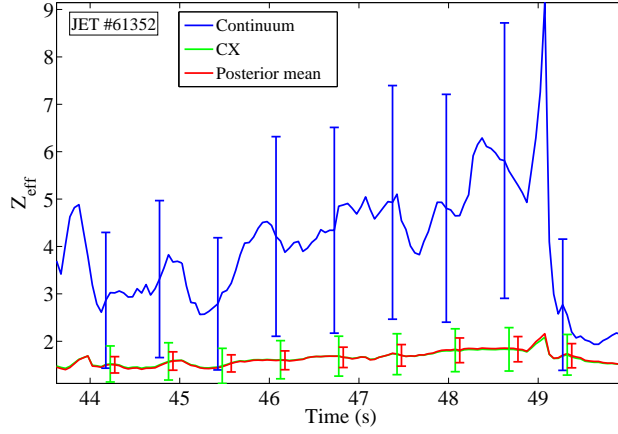


Figure 6.17: Marginal posterior means of the on-axis Z_{eff} for JET #61352, including also LIDAR measurements with an increased error of 20%.

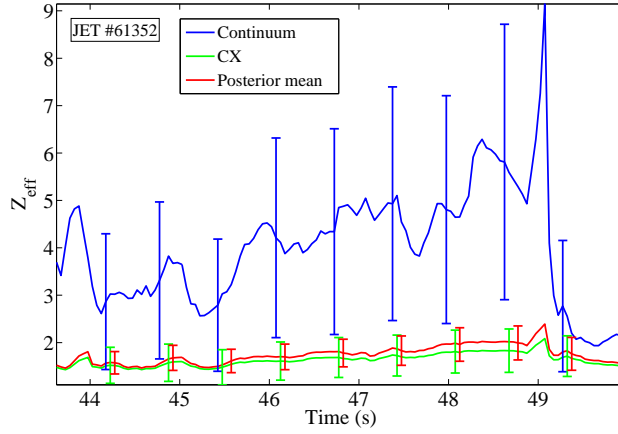


Figure 6.18: Marginal posterior means of the on-axis Z_{eff} for JET #61352, including also LIDAR measurements with an increased error of 20%. The error on δ has also been increased to 50%.

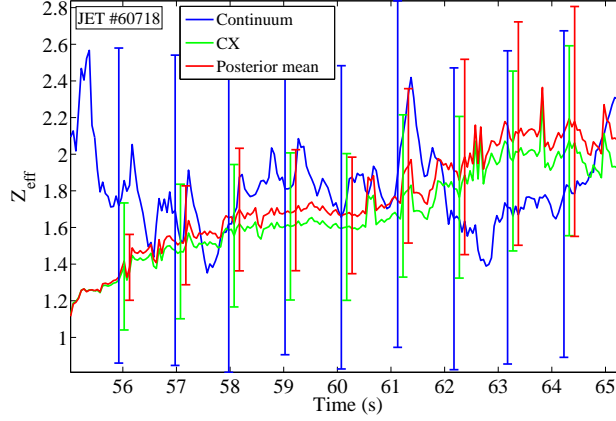


Figure 6.19: Marginal posterior means of the on-axis Z_{eff} for JET #60718, including also LIDAR measurements with an increased error of 20%. The error on δ has also been increased to 50%.

a pure continuum likelihood, a pure CX likelihood, a combination of continuum and LIDAR, a combination of CX and LIDAR, a combination of continuum and CX, and finally a combination of all three likelihood factors. Since the pure LIDAR posterior reaches high values, we have made a zoomed display of this figure in Figure 6.21. The posterior densities for Z_{eff} are shown in Figure 6.22. The marginals for n_e and Z_{eff} for the combination of all three likelihoods have a very low density, so they are shown separately in Figures 6.23 and 6.24. These two marginals are symmetric, bell-shaped curves, so their maximum corresponds to the value of the estimated n_e , respectively Z_{eff} at 47 s, displayed in Figures 6.13 and 6.14. Note how the probability densities are all cut off at the boundaries of the respective uniform priors (except of course the distribution for Z_{eff} given only the LIDAR measurements, since the LIDAR likelihood does not depend on Z_{eff}).

The most striking features of this overview of marginal posteriors are the following. From Figure 6.21, it can be seen that both the CX (δ) and especially the continuum data (ϵ) are consistent with the LIDAR measurements. This can be seen through the fairly large overlap of the respective probability distributions. The density estimated from the combination of continuum and CX likelihoods (i.e. approximately the maximum of the black curve), is considerably higher than the LIDAR density. This is a result already obtained in Figure 6.10. Quite surpris-

ingly, for the Z_{eff} marginals in Figure 6.22, the continuum and CX data are actually very well consistent—albeit over a broad range—resulting in the black curve. So, the Z_{eff} estimate given in Figure 6.11 is consistent with each of the continuum and CX datasets alone. However, since the accompanying n_e estimate differed so much from the LIDAR density, we chose to reject also the Z_{eff} estimate. It is only as soon as also the LIDAR data are added to the continuum data on the one hand (magenta curve), and the CX data on the other hand (cyan curve), that the corresponding estimates for Z_{eff} become inconsistent. These are exactly the traditional estimates one becomes by calculating Z_{eff} from, respectively, ϵ_{ff} and $n_{e,\text{I}}$ on the one hand, and δ and $n_{e,\text{I}}$ on the other hand. Finally, if one tries to combine the ϵ , δ and $n_{e,\text{I}}$ measurements, one arrives at a marginal posterior with a very low density, Figure 6.24, which is more or less consistent with the CX Z_{eff} , but not at all with the continuum Z_{eff} . From the above analysis it also follows that the data for the other discharge under study, #60718, will be much more self-consistent than for #61352.

Thus, we see that the high trust we put in the LIDAR density measurements seems at first sight to result in inconsistency of the total data set. It is clear that the consistency of the data would be better if the uncertainty on the LIDAR measurements were allowed to be higher, resulting in a broader LIDAR posterior. As we have seen in the previous subsection, this would have the effect of the Z_{eff} estimate to be the most consistent with the CX Z_{eff} , as compared to the continuum Z_{eff} . However, at this stage we should not jump to the conclusion that there is a higher degree of uncertainty in the $n_{e,\text{I}}$ measurements than expected before. It is, indeed, still possible that a more sophisticated BIDA scheme, where more uncertainties are modelled, or where the uncertainties are modelled in more detail, will reveal other sources of error that are responsible for the here observed data inconsistency. In the next subsection, we will see that this is actually the case.

Indeed, it will become clear that in general one should not neglect the effect of *systematic* uncertainties on the input data. Thus, it is true that we have succeeded in calculating a Z_{eff} that is consistent with both input data sets, but this should always be seen in the context of the proposed model. If we do not model all sources of uncertainty to a sufficient extent, we should not expect that our estimated Z_{eff} , although consistent with all input data, brings us much closer to the real, physical Z_{eff} value. Another indication to a possible lack in the

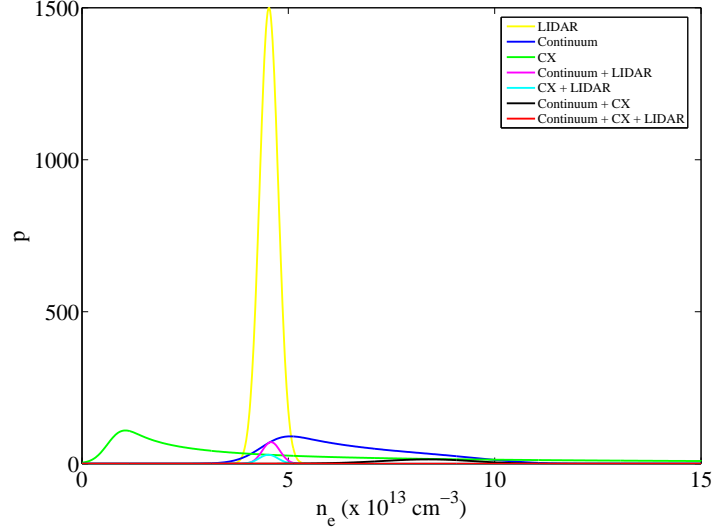


Figure 6.20: Marginal posterior distributions for the on-axis n_e for JET pulse #61352 at 47s, taking into account different data sets.

modelling of uncertainties is that the posterior density for both n_e and Z_{eff} , taking into account all data (including LIDAR), is very low compared to the other densities in Figures 6.21 and 6.22.

Still, as we will see in the next subsection, the analysis taking into account only statistical uncertainties can be valuable to retrieve the relative time evolution of Z_{eff} . However, in cases where systematic uncertainties are important, the method does not permit to find the correct absolute Z_{eff} value.

6.4.3 Inference from the local model: statistical and systematic uncertainties

6.4.3.1 Model definition

In the previous section we have seen that, by modelling the statistical uncertainties in the data, it was possible to arrive at a Z_{eff} estimate consistent with both the continuum and CX Z_{eff} . However, for the JET discharge #61352, there is a high, systematic data inconsistency. We now go to the next level of modelling also systematic uncertainties on the ϵ and δ measurements, and the question is whether this will

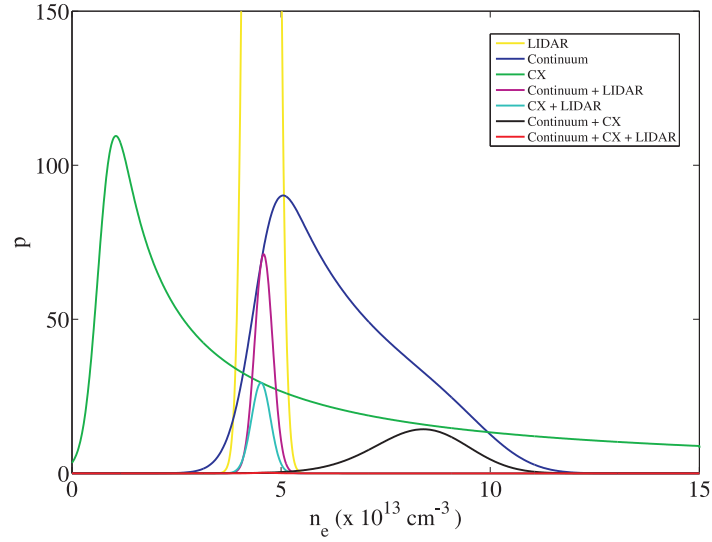
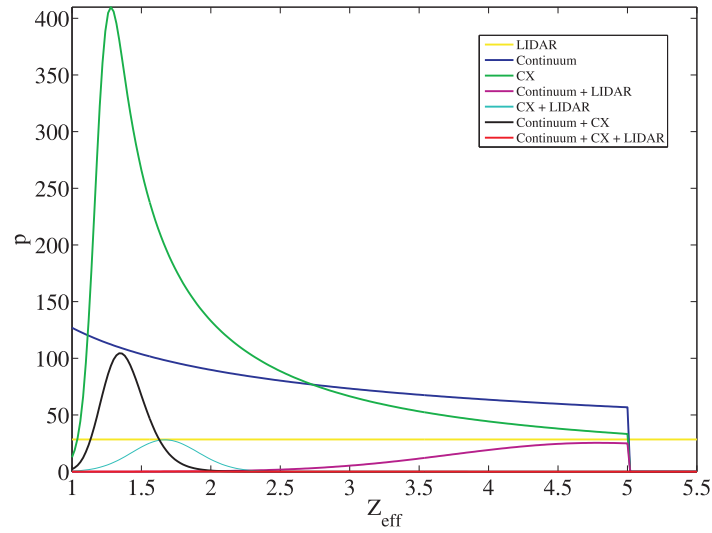


Figure 6.21: Zoomed display of Figure 6.21.

Figure 6.22: Marginal posterior distributions for the on-axis Z_{eff} for JET pulse #61352 at 47s, taking into account different data sets.

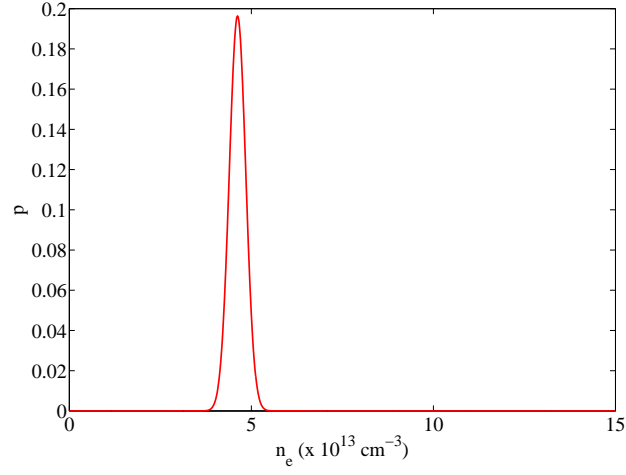


Figure 6.23: The marginal posterior for the on-axis n_e for the combination of all three likelihoods.

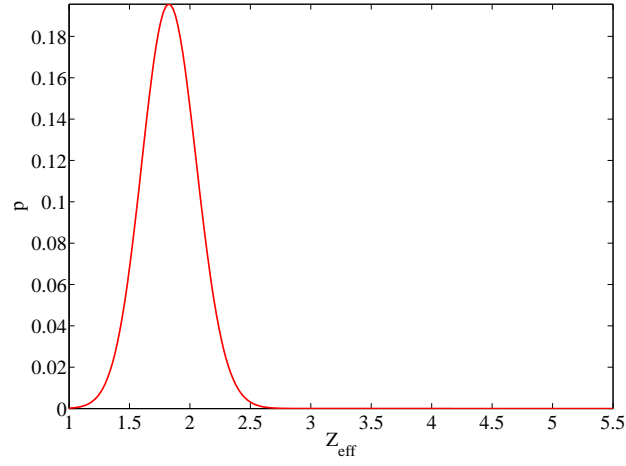


Figure 6.24: The marginal posterior for the on-axis Z_{eff} for the combination of all three likelihoods.

increase the overall data consistency. We will also explicitly incorporate LIDAR density measurements, but we will not consider any systematic uncertainty on the LIDAR density. The statistical uncertainties on the input data ϵ and δ have been taken of the same magnitude as in the previous subsection. In any case, the results do not depend very much on the exact value of the statistical uncertainties.

Incidentally, we would like to mention that an often made (silent) mistake in experimentation is precisely the consideration of errors on a measurement as being of a pure statistical nature. This can sometimes be noticed implicitly in a case where the experimentalist applies simple Gaussian error propagation laws.

As a first approximation, we will assume that the measured data ϵ and δ may be off the real physical value by a scale factor. This scale factor may or may not be of interest, thus determining its status as a parameter of interest, or rather as a nuisance parameter, respectively. The full posterior density, taking into account LIDAR measurements as well, becomes:

$$p(n_e, Z_{\text{eff}}, s_\epsilon, s_\delta | \epsilon, \delta, n_{e,I}, I) \sim \exp \left[-\frac{\gamma_1}{2} (\epsilon - s_\epsilon n_e^2 Z_{\text{eff}})^2 - \frac{\gamma_2}{2} (\delta - s_\delta n_e [Z_{\text{eff}} - 1])^2 - \frac{\gamma_I}{2} (n_{e,I} - n_e)^2 \right] U_{n_e} U_{Z_{\text{eff}}} U_{s_\epsilon} U_{s_\delta}. \quad (6.12)$$

There are now four unknown parameters, where s_ϵ stands for the scale factor belonging to the measurement ϵ , and likewise for δ . The prior distributions for the scale factors are relatively important. Indeed, since we dispose of only three measurements, the prior should be sufficiently uninformative, see Section 5.1.6. Several (relatively) uninformative distributions for the scale parameters were tested, but eventually it was decided to assume simple uniform distributions, defined by the following boundaries, which we assume to be reasonable:

$$0.2 < s_\epsilon, s_\delta < 1.7.$$

6.4.3.2 Metropolis sampling

In principle, to find posterior mean values for n_e , Z_{eff} , s_ϵ and s_δ , we could proceed in a similar way as in Section 6.4.2, thus calculating the joint posterior density on a grid of values for the parameters, and then calculating marginal means. However, with four parameters this would already become quite time-consuming. The analytical calculation of

the multidimensional integrals is not possible, especially because of the appearance of both n_e and n_e^2 in the joint posterior. It is on the other hand possible to use gradient methods to search for the maximum of the joint posterior, and thus the marginal maxima, but, depending on the skewness of the distribution, the maxima could differ substantially from the marginal means. Instead of utilizing some numerical integration method, we choose to sample from the marginal distributions using a Markov Chain Monte Carlo algorithm. Numerical integration methods become very time-consuming anyway, when the dimensionality of the problem increases. Moreover, the MCMC scheme developed here will be of great help in the next section, where we will directly estimate Z_{eff} profiles. Finally, for the consistency analysis in Section 6.4.3.3, we will want to display the marginal distributions for the parameters of interest, so we need in any case a sampling method.

We will use the Metropolis-Hastings algorithm, since the posterior is intractable as a function of n_e . As a proposal distribution, we employ the Cauchy distribution, in physics also known as the Lorentz distribution. It has the advantage of being rather peaked on the one hand, but still having broad tails, so that the probability of sampling far away from its mode (i.e. the current value in the Markov chain) is sufficiently high. This ensures that the support of the posterior is scanned effectively. The Cauchy distribution for θ with location parameter x_0 and scale parameter κ has the following probability density:

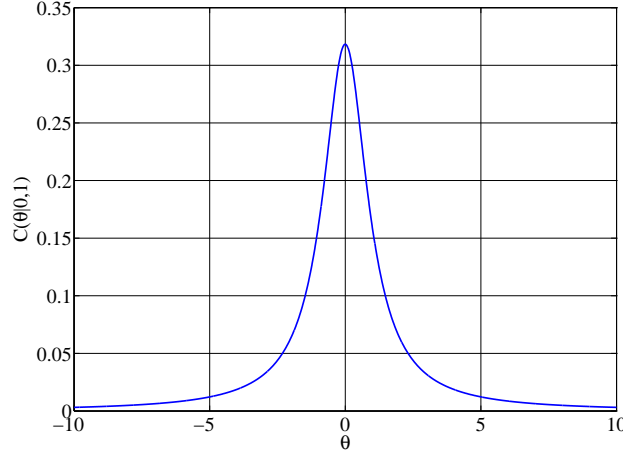
$$C(\theta|x_0, \kappa) = \frac{1}{\pi} \frac{\kappa}{(x - x_0)^2 + \kappa^2}, \quad \kappa > 0.$$

As an example, the Cauchy distribution $C(\theta|0, 1)$ (i.e. the standard Cauchy distribution) is plotted in Figure 6.25. Since the Cauchy distribution is symmetric in its argument versus its location parameter, it drops out of the acceptance probability, and we effectively use the Metropolis algorithm. The Cauchy inverse cumulative distribution function for a probability p is given by

$$F^{-1}(p|x_0, \kappa) = x_0 + \kappa \tan[\pi(p - 1/2)],$$

so that the Markov chain is a random walk chain, by the definition given in Section 5.2.4.1.

An MCMC code was written in the frame of the current work. The Markov chain was run in parallel for all time points of the measured time traces for ϵ and δ . The starting values for the entire n_e time series

Figure 6.25: The standard Cauchy distribution $C(\theta|0, 1)$.

were taken at $2 \times 10^{13} \text{ cm}^{-3}$, and for the entire Z_{eff} time series at 2. The scale parameters s_ϵ and s_δ were initialized at unity. The Cauchy scale parameters κ were chosen so as to maintain a post-convergence acceptance probability in the range 40% – 60%. This is a level recommended by Roberts, Gelman and Gilks [111] based on empirical studies, providing good convergence properties of the chain. The burn-in period was chosen to be 4000 samples, while the Monte Carlo period was 2000 samples long. To check the code for bugs, first a four-dimensional uniform distribution was taken as a target distribution. This should result in a uniform marginal distribution for all parameters. Data from JET discharge #61352 were then used to construct a histogram from the Z_{eff} samples at 47s, shown in Figure 6.26. In order to be able to construct a sufficiently complete histogram, the number of Monte Carlo samples was temporarily increased to 30.000. It is clear that the uniform distribution was sampled correctly. Similar histograms can be constructed for the other parameters.

Artificial data set As the flexibility of our model has increased with the addition of the scale parameters, we shall first run the estimation algorithm on an artificial data set. We start from an artificial n_e and Z_{eff} time trace, which we assume to be the real physical n_e and Z_{eff} . As an example, we take the measured time traces of n_e from LIDAR

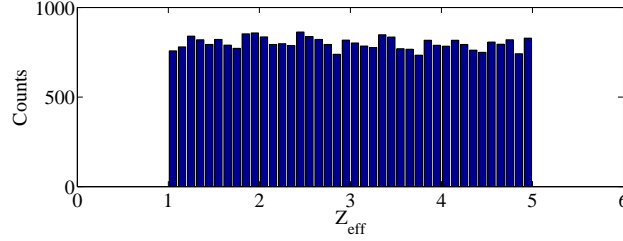


Figure 6.26: Histogram of the sampled marginal distribution of the on-axis Z_{eff} after a Monte Carlo run of 30,000 samples, using an artificially imposed uniform posterior distribution ($1 \leq Z_{\text{eff}} \leq 5$).

and Z_{eff} from bremsstrahlung in JET pulse #60718. The corresponding times traces of ϵ and δ were calculated, and artificial scale factors were introduced: ϵ was multiplied by a factor $s_\epsilon = 1.2$, while δ was multiplied by a factor $s_\delta = 0.3$. This simulates a small systematic overestimation of ϵ , and a substantial underestimation of δ , and thus an underestimation by the same factor s_δ of the summed impurity density. The results of the estimation for the artificial data set are shown in Figures 6.27 and 6.28. The density has been estimated almost precisely, and the estimate for Z_{eff} is also very good. The estimate for s_ϵ was precisely 1.2, and for s_δ 0.31.

A word about the error bars is in place. The error bars shown for the traditional n_e and Z_{eff} values, are still the error bars quoted in Table 4.2, on the understanding that they should now be viewed as containing both statistical and systematic effects. As far as the error bars on the MCMC estimates are concerned, since we have taken into account the systematic uncertainties already, we assume that the remaining uncertainty is of a pure statistical nature. In the same way as we calculate Monte Carlo estimates of the marginal mean of the parameters of interest, we can calculate estimates for the standard deviation σ . The standard deviation of the Monte Carlo mean is then given by $\sigma/\sqrt{n_{\text{MC}}}$, where n_{MC} is the number of Monte Carlo samples (i.e. post-convergence samples). However, this is only true when all samples are statistically independent, whereas we know that the Markov chain generates samples that are to a certain extent correlated. On the other hand, it is a common feature of a Markov chain that the farther two states of the chain are away from each other in ‘time’, the more they be-

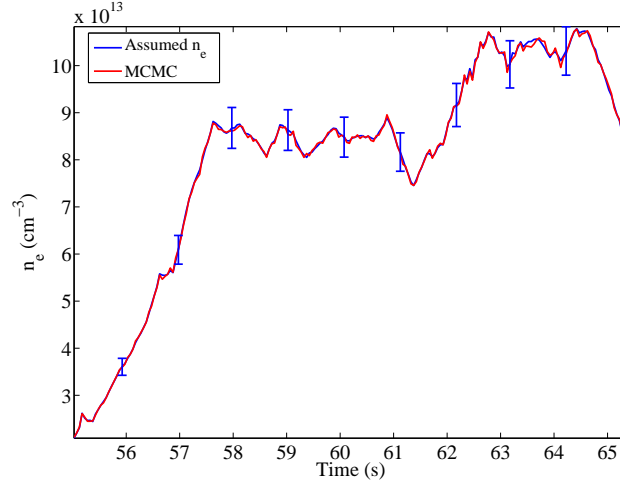


Figure 6.27: Monte Carlo mean of the on-axis n_e using an artificial data set.

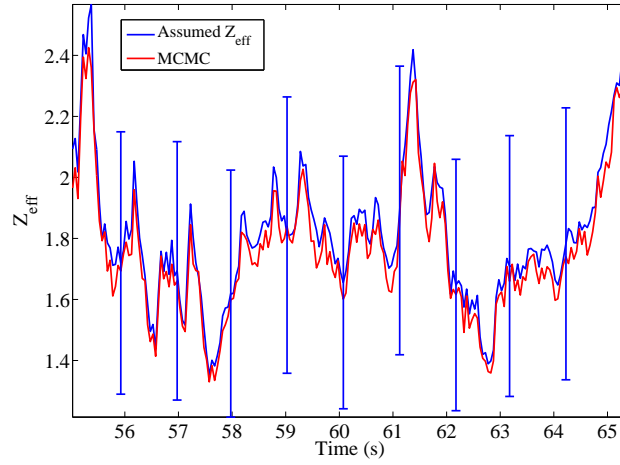


Figure 6.28: Monte Carlo mean of the on-axis Z_{eff} using an artificial data set.

come decorrelated. The correlation between two states $X^{(t)}$ and $X^{(t+s)}$ is measured by the autocorrelation function

$$r(s) = \frac{1}{\sigma^2} \text{E}[(X^{(t)} - \langle X \rangle)(X^{(t+s)} - \langle X \rangle)],$$

where s is called the *lag*. We then define

$$\tau = 1 + 2 \sum_{s=1}^{\infty} r(s)$$

as the number of dependent states from the Markov chain that are the equivalent of one independent state [112]. It is often observed that $r(s)$ is roughly exponential in shape; that is, the correlation between two states is falling exponentially with the lag between the states. τ can then also be seen as the time constant governing this exponential. The smaller τ , the more efficient the MCMC sampler is. The autocorrelation as a function of the lag s for a typical MCMC chain for n_e (see below), has been plotted in Figure 6.29. An arbitrary state $X^{(t)}$ within the Monte Carlo phase of the chain was used as a reference, and the autocorrelation was calculated between this reference state, and a state $X^{(t+s)}$, $s = 0, \dots, 19$, some time lag s away from the reference. Obviously, the correlation between the reference state and itself is 1, while the correlation drops to practically zero at $s = 19$. The calculated τ in this case is 7.6, meaning that approximately eight samples of the chain are the equivalent of one independent sample. The decay occurs roughly exponentially, and τ can indeed be seen as an approximate time constant of this decay. Now, since n_{MC}/τ is an estimate for the total number of *independent* Monte Carlo samples, we have to use the following expression for the standard deviation σ_{MC} on the Monte Carlo mean:

$$\sigma_{\text{MC}} = \sigma / \sqrt{n_{\text{MC}}/\tau} = \sqrt{\left[\frac{1}{n_{\text{MC}} - 1} \sum_t (x^{(t)} - \langle X \rangle)^2 \right] / \left(\frac{n_{\text{MC}}}{\tau} \right)},$$

where the sum is over all Monte Carlo (post-convergence) samples. Thus, the effect of the dependence of the Monte Carlo samples is to increase the standard deviation on the Monte Carlo mean $\langle X \rangle$. The thus calculated error bars for the MCMC estimates are not displayed in Figures 6.27 and 6.28 since they are very small (less than one percent). This indicates that the MCMC n_e and Z_{eff} are very good estimates of the respective marginal means. Also in the remainder of the analysis we will not show the small MCMC error bars.

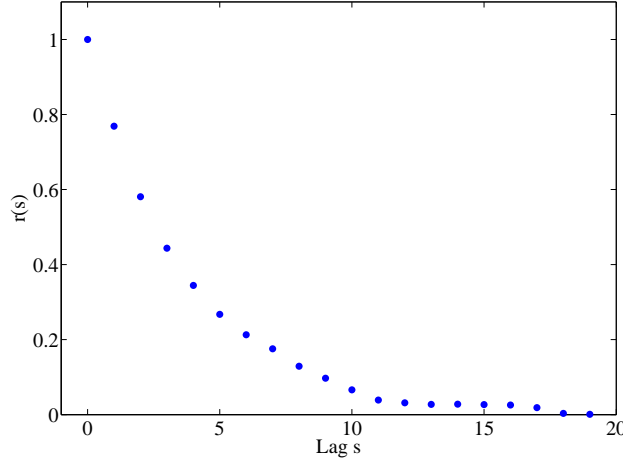


Figure 6.29: Autocorrelation between n_e samples in an MCMC chain as a function of time lag s .

Real data set Next, we performed the MCMC estimation based on real measured continuum and CX data sets. For pulse #60718, the results of an MCMC run with a Monte Carlo period of 2000 samples are displayed in Figures 6.30 to 6.33. Figure 6.30 shows the time traces for all samples of the sampled n_e and Z_{eff} at 63 s. Note the excursion of both the n_e and Z_{eff} traces from an initial value to a steady state. This is the effective burn-in period of the chain, where it is searching for convergence. Figure 6.31 contains the trace of the joint posterior probability density. Finally, Figures 6.32 and 6.33 show the results of the estimation. The estimated density is a very good estimate of the LIDAR measurement, indicating that the numerical calculation is sound. The behaviour in time of the estimated Z_{eff} is very similar to the evolution of the CX Z_{eff} . It is rather surprising that the estimate lies for most of the time slightly below both the continuum and CX Z_{eff} . The estimated s_ϵ equals 1.24 ± 0.01 while $s_\delta = 1.43 \pm 0.03$ ⁴. The values for the scale parameters indicate that the continuum, CX and LIDAR data can be made consistent if the ϵ data are decreased with a factor $1/s_\epsilon = 0.81$ and the δ data with a factor $1/s_\delta = 0.70$.

Next, data were used from pulse #61352, using the same parameters as for the #60718 MCMC run, except for the Cauchy scales, which were

⁴The errors on s_ϵ and s_δ are calculated in a similar way as for n_e and Z_{eff} .

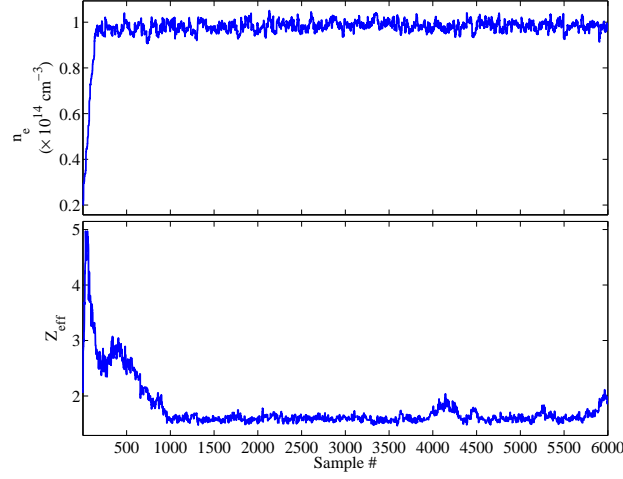


Figure 6.30: MCMC trace of the sampled on-axis n_e and Z_{eff} at 63 s in JET #60718. There is a clear effective burn-in of about 1000 samples.

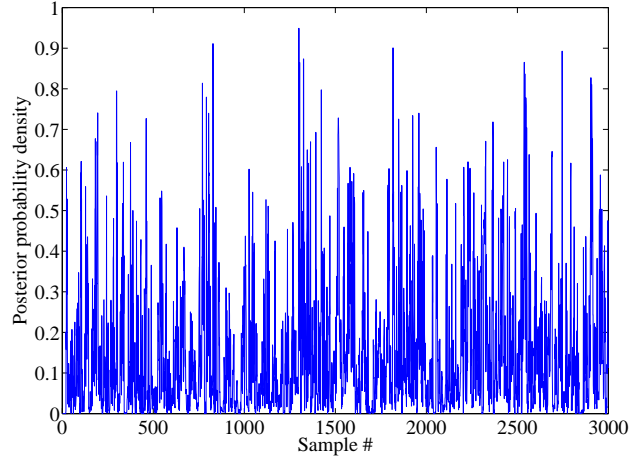


Figure 6.31: Joint posterior probability density in an MCMC run using data from JET #60718.

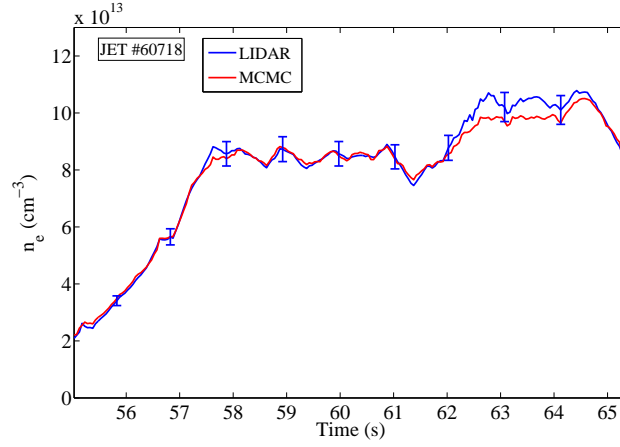


Figure 6.32: Monte Carlo mean of the on-axis n_e using data from JET #60718.

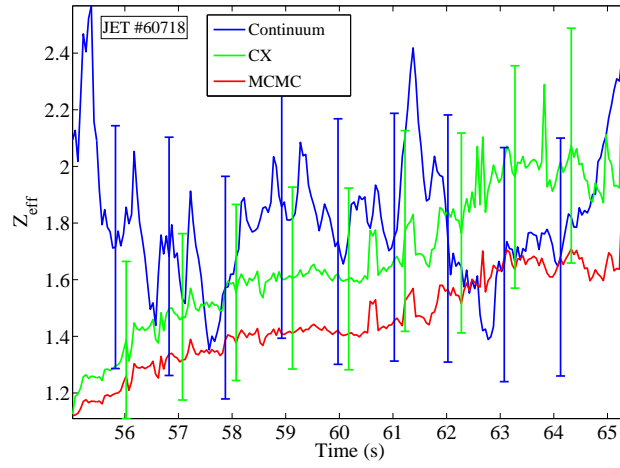


Figure 6.33: Monte Carlo mean of the on-axis Z_{eff} using data from JET #60718

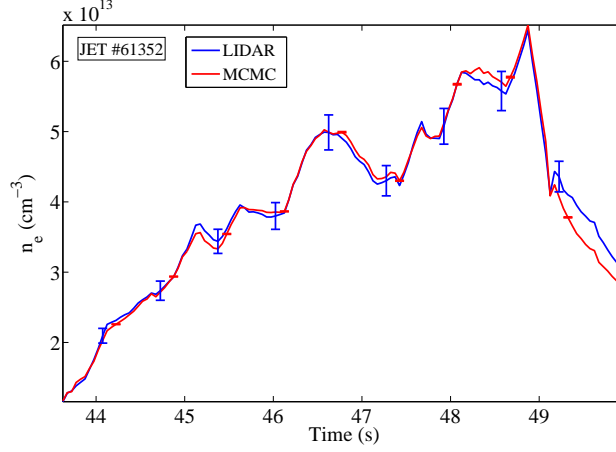


Figure 6.34: Monte Carlo mean of the on-axis n_e using data from JET #61352.

tuned so as to obtain favourable acceptance rates. The estimation results of the run for n_e and Z_{eff} are shown in Figures 6.34 and 6.35. Again, the sampled density is very close to the density from LIDAR. The sampled Z_{eff} now lies in between the continuum Z_{eff} and the CX Z_{eff} , and the estimate is in considerable contrast to the estimates given in Section 6.4.2, where only statistical errors on the data were taken into account. The estimated scale parameter s_ϵ is 1.52 ± 0.02 , and $s_\delta = 0.40 \pm 0.01$. The values for the scale parameters indicate that the continuum, CX and LIDAR data can be made consistent if the ϵ data are decreased by a factor $1/s_\epsilon = 0.66$ and the δ data are increased with a factor $1/s_\delta = 2.50$.

We have tested the method on a few more data sets from JET discharges. The results for JET pulse #59193 are shown in Figures 6.36 and 6.37. The Z_{eff} estimate is, perhaps surprisingly, quite close to the CX Z_{eff} , rather than to the continuum Z_{eff} . The estimates for the scale parameters are $s_\epsilon = 1.45 \pm 0.01$ and $s_\delta = 0.80 \pm 0.02$.

Next come the estimation results for JET #59186, in Figures 6.38 and 6.39. The Z_{eff} estimate is again close to the CX Z_{eff} , rather than to the continuum Z_{eff} . The estimates for the scale parameters are $s_\epsilon = 1.30 \pm 0.01$ and $s_\delta = 0.84 \pm 0.03$; almost the same as for #59193.

We finally display the estimation results for JET #59194, in Figures 6.40 and 6.41. The Z_{eff} estimate is again more similar to the CX

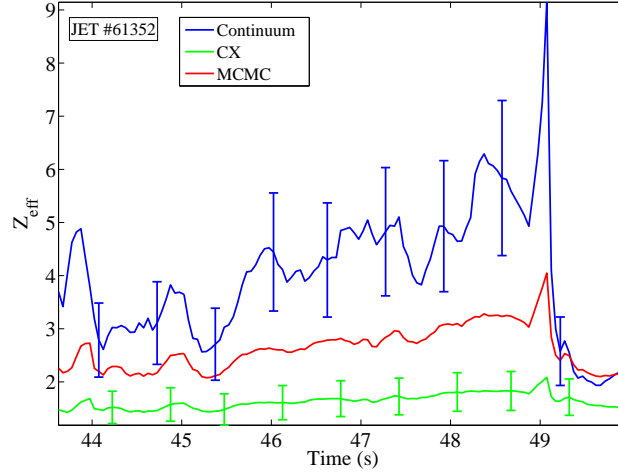


Figure 6.35: Monte Carlo mean of the on-axis Z_{eff} using data from JET #61352.

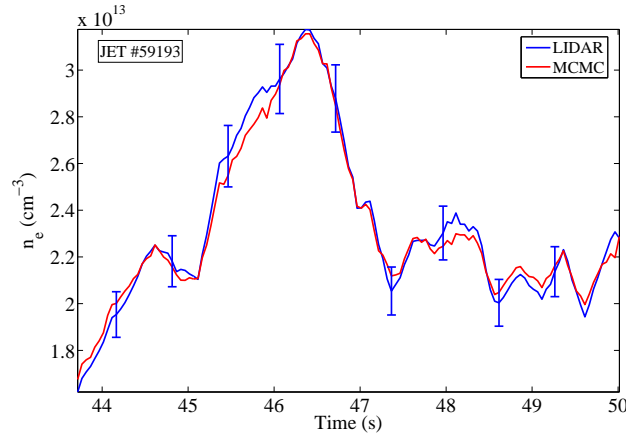


Figure 6.36: Monte Carlo mean of the on-axis n_e using data from JET #59193.

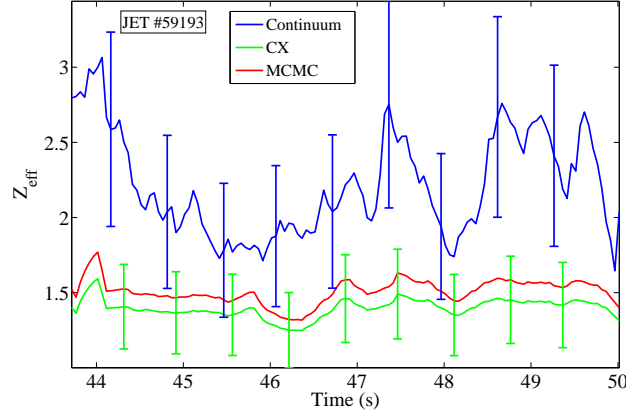


Figure 6.37: Monte Carlo mean of the on-axis Z_{eff} using data from JET #59193.

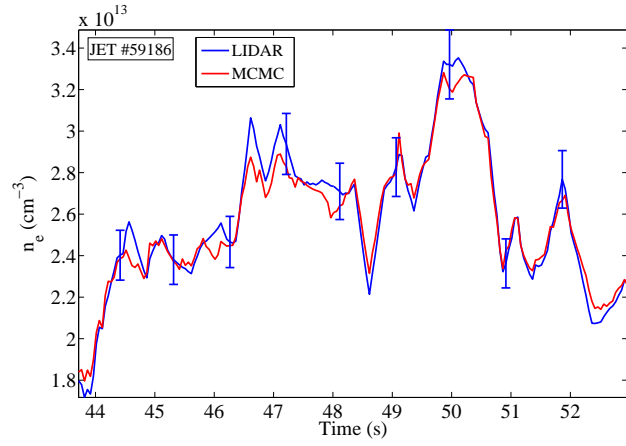


Figure 6.38: Monte Carlo mean of the on-axis n_e using data from JET #59186.

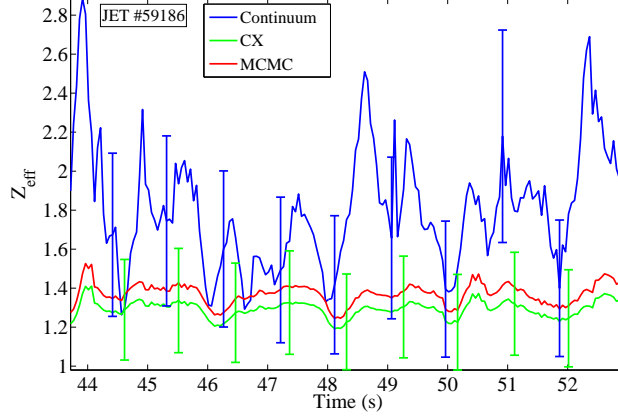


Figure 6.39: Monte Carlo mean of the on-axis Z_{eff} using data from JET #59186.

Z_{eff} , rather than to the continuum Z_{eff} . The estimates for the scale parameters are $s_{\epsilon} = 1.44 \pm 0.01$ and $s_{\delta} = 1.03 \pm 0.03$.

6.4.3.3 Consistency analysis

We can visualize the increased consistency between the scale corrected measurements ϵ and δ , by again performing a consistency analysis for pulse #61352 at 47 s. In calculating the joint posterior densities, we have taken scale factors into account as estimated before, namely $s_{\epsilon} = 1.52$ and $s_{\delta} = 0.40$. The results of the analysis are displayed in Figure 6.42 for the density marginals (zoomed display), and Figure 6.43 for the Z_{eff} marginals. It is apparent that all marginals overlap for a large part, indicating a considerable increase of consistency between the scale corrected data ϵ and δ , compared to the case in Section 6.4.2.3, where only statistical data uncertainties were taken into account.

6.4.3.4 Conclusion

We can conclude from this section that adding the extra freedom of a scale parameter for the measured data leads to Z_{eff} estimates that are more consistent with both the continuum and CX Z_{eff} . In addition, the scale parameters allow to render the continuum and CX input data more consistent. The inclusion of the scale parameters increases the sophistication and flexibility of the Bayesian probabilistic model, since

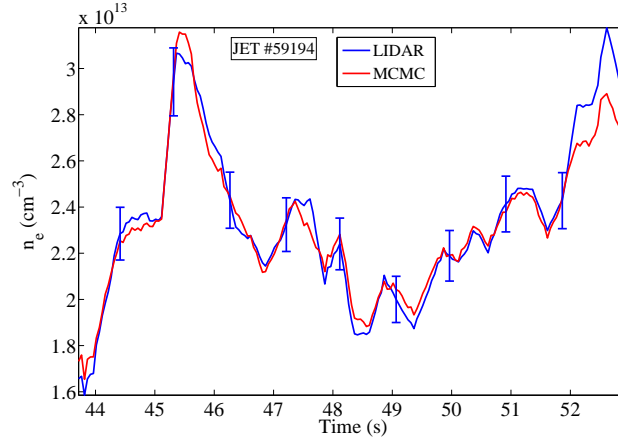


Figure 6.40: Monte Carlo mean of the on-axis n_e using data from JET #59194.

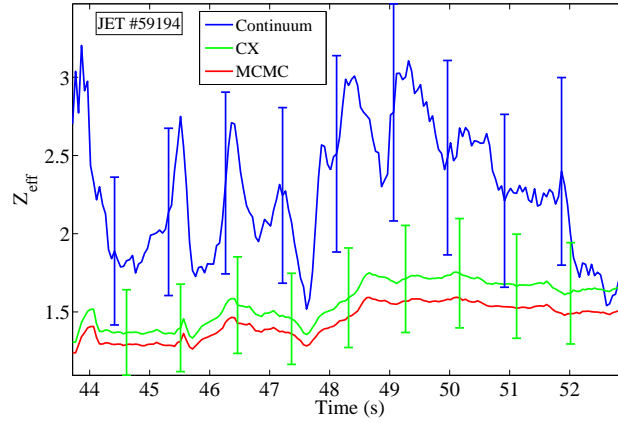


Figure 6.41: Monte Carlo mean of the on-axis Z_{eff} using data from JET #59194.

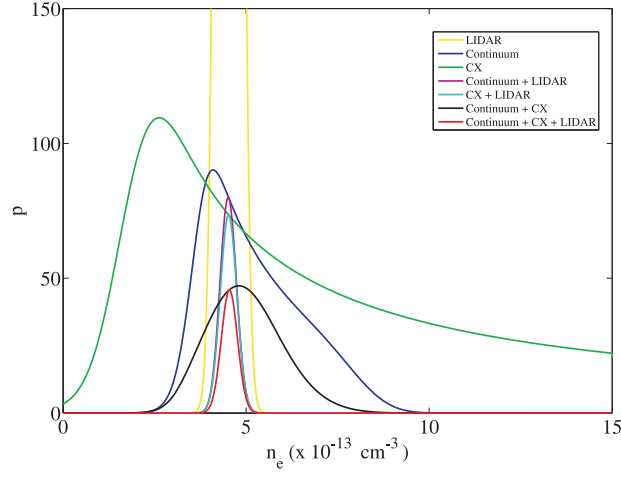


Figure 6.42: Marginal posterior distributions for the on-axis n_e for JET pulse #61352 at 47s, taking into account different data sets. The ϵ and δ measurements have been scale corrected.

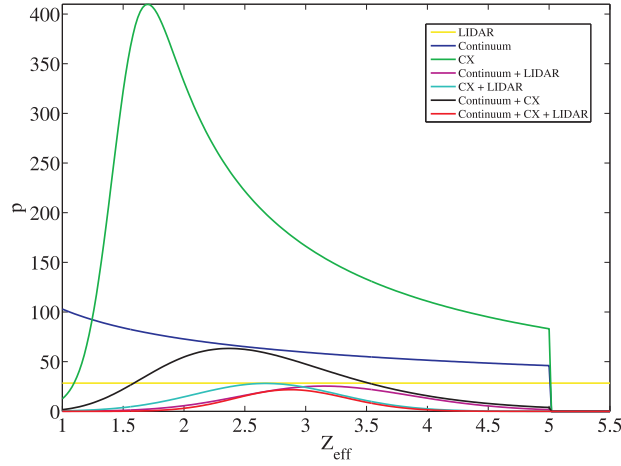


Figure 6.43: Marginal posterior distributions for the on-axis Z_{eff} for JET pulse #61352 at 47s, taking into account different data sets. The ϵ and δ measurements have been scale corrected.

now also the systematic uncertainties are modelled. Therefore, in cases where the systematic discrepancy between the continuum and CX Z_{eff} is considerable, we choose to attach more value to the Z_{eff} traces estimated in this section, compared to the case where only the statistical uncertainties were modelled. Moreover, the amount of consistency of the scaled input data has been improved to such a degree, that we feel that we might, even with the still rather simple model, already put good trust in the obtained Z_{eff} estimates. Indeed, we now dispose of two entirely different diagnostic techniques whose measurements are mutually consistent. The enhanced Z_{eff} accuracy is also reflected by the error bars, which have shrunk dramatically compared to the error bars on the continuum and CX Z_{eff} . The increased data consistency also indicates that the proposed model includes a relatively good description of the various uncertainties. Nevertheless, it would still be valuable to obtain more detailed information on the systematic uncertainties on ϵ and δ , which can then be encoded in the prior distributions for s_ϵ and s_δ . It is also expected that, by still increasing the model sophistication, and by modelling gradually more sources of uncertainty, the data consistency and the quality of the Z_{eff} estimates will still increase.

An overview of the estimated scale factors for the various discharges under study is given in Table 6.2. In all discharges under study, the continuum ϵ appears to be seriously overestimated, leading to an overestimation of the continuum Z_{eff} . Since the influence of electron temperature on ϵ is small, this can most likely be related to an overestimation of the bremsstrahlung emissivity. We have estimated a factor by which ϵ should be multiplied, which is on average 0.72. A possible cause for overestimation might be the contribution of edge non-bremsstrahlung components in the continuum, which propagate towards the centre during profile reconstruction. Residual in-vessel reflections and the occurrence of atomic lines in the observed spectral window might be other reasons. For discharge #61352, the ϵ overestimation can be linked to the unphysically high values of over 6 of the continuum Z_{eff} .

On the other hand, in three of the considered discharges, namely #61352, #59193 and #59186, the δ data (the summed impurity density) seem to be underestimated, yielding an underestimation of the CX Z_{eff} . δ should be multiplied with a factor of 2.5, 1.25 and 1.19, respectively. If we assume that only fully stripped carbon makes a contribution to the on-axis Z_{eff} , this means that the CX carbon density should be multiplied by the same factor. If, on the other hand, one

Discharge #	s_ϵ	s_δ
60718	1.24 ± 0.01	1.43 ± 0.03
61352	1.52 ± 0.02	0.40 ± 0.01
59193	1.45 ± 0.01	0.80 ± 0.02
59186	1.30 ± 0.01	0.84 ± 0.03
59194	1.44 ± 0.01	1.03 ± 0.03

Table 6.2: MCMC estimated scale factors for the on-axis ϵ and δ in several JET discharges.

suspects that also other impurities make a non-negligible contribution to Z_{eff} , this adds corresponding terms in the summation (6.1). Since we have only modelled a scale factor for δ (and no offset), we should then assume, if we want this model to hold, that the time evolution of all impurity densities considered, is very similar. Indeed, the densities of the additional impurities should be related to the carbon density by a simple factor. This is usually the case, and we could then, for example, understand the underestimation of δ in terms of the influence of not monitored impurities. Otherwise, also an unknown offset term to delta, variable in time, should be taken into account. We will not consider this at this point.

Since in the pulses #59193, #59186 and #59194, the estimated Z_{eff} is more similar both qualitatively and quantitatively, to the CX Z_{eff} than to the continuum Z_{eff} , for these pulses we are tempted to put more trust in the CX Z_{eff} than in the continuum Z_{eff} . This is especially true for discharge #59194, which has an s_δ scale factor very close to unity.

Finally, note that we have assumed that, throughout a single discharge, s_ϵ and s_δ remain constant, a hypothesis that might also be changed in the future.

6.5 IDA for the estimation of Z_{eff} profiles

The analysis in the previous section can in principle be carried out at different radial positions as well, not only on the magnetic axis. While this approach can indeed yield valuable information about the Z_{eff} profile that is consistent with both the continuum and CX data sets, we wish to explore a different route in this section. After all, eventually

we want to formulate a BIDA model that allows us to directly make inferences about the parameters of interest, starting from the raw, uncalibrated measurements of the respective detectors. Then, uncertainties on the calibration, together with all other important sources of uncertainty, can be properly modelled in terms of likelihood distributions. In a next step towards the realization of this goal, we wish to start directly from the line-integrated, but calibrated, bremsstrahlung emissivity measurements and line-integrated density measurements, together with local data on impurity densities from CX. From this, we wish to derive radial Z_{eff} profiles. This strategy has the great advantages that we do not need an external inversion procedure, and that we can model uncertainties on the magnetic equilibrium into the Bayesian probability model. The line intensities measured by CXS are already local plasma properties, since they originate from the plasma volume intersected by the neutral beam. We will first outline the principle of the method, and later perform some tests on an artificial data set. This method has not yet been applied on a real data set.

6.5.1 Model definition

6.5.1.1 Geometry matrix

Our approach to the problem is very similar to the one for the local model. First, we formulate a likelihood for the parameters of interest. To this end, we consider a poloidal plasma cross-section viewed by a set of chords belonging to the bremsstrahlung diagnostic and a set of the CX diagnostic. We assume that the line-integrated electron density data originate from interferometry. These line-integrals are obtained by observing the interference patterns between a reference laser beam, and a set of beams traversing the plasma, which also define a set of chords. For simplicity, a configuration with circular concentric magnetic surfaces has been taken as an example in Figure 6.44. We denote the radial n_e and Z_{eff} profiles by $n_e(r)$ and $Z_{\text{eff}}(r)$, while sometimes we will use the shorthand notations n_e and Z_{eff} . r should be understood as a general notation for a ‘radial’ parameter in some coordinate system. It could be, for example, the radius of magnetic surfaces measured along the midplane. We assume that all plasma parameters are constant on magnetic flux surfaces. For numerical reasons, all profiles

$\eta(r)$ are approximated by a piecewise constant function:

$$\eta(r) = \sum_{j=1}^{n_r} \eta_j \chi_j(r), \quad (6.13)$$

where $\chi_j(r)$ denotes the indicator function on the interval $(j-1)\frac{a}{n_r} \leq r < j\frac{a}{n_r}$, which equals unity only in this interval and zero otherwise. The LCFS is assumed to be labeled by $r = a$, and n_r denotes the number of considered flux surfaces, beginning at the magnetic axis for $j = 1$. Naturally, the higher n_r , the more precise the approximation (6.13). Again for simplicity, we suppose that the lines of sight (or laser beams) are parallel to the X axis in Figure 6.44, with intercepts on the Y axis given by

$$y_i = \left(i - \frac{1}{2}\right) \frac{a}{n_c}, \quad i = 1, \dots, n_c,$$

with n_c the number of channels for the corresponding diagnostic. If we now set

$$r_j = j \frac{a}{n_r}, \quad (6.14)$$

we can define a matrix \widetilde{M} given by

$$\widetilde{M}_{ij} = \begin{cases} 2 \left(\sqrt{r_j^2 - y_i^2} - \sqrt{r_{j-1}^2 - y_i^2} \right) & \text{if } y_i \leq r_{j-1} \\ 2 \sqrt{r_j^2 - y_i^2} & \text{if } r_{j-1} < y_i \leq r_j \\ 0 & \text{otherwise} \end{cases}.$$

This is the geometry matrix, which we already encountered before in Section 4.1.3, allowing us to approximate the line-integral along the line of sight i over the profile $\eta(r)$ as

$$L_{\eta,i} = \sum_{j=1}^{n_r} \widetilde{M}_{ij} \eta_j. \quad (6.15)$$

Again, the higher n_r , the more accurate this approximation is. The idea behind an inversion method for the reconstruction of profiles from line-integrated data, is to invert the matrix equations (6.15). This results in an estimate for the parameters η_j . Many different approaches can then be applied to ensure that the resulting profiles are sufficiently regular, since the geometry matrix typically is close to singular. Some of these methods were highlighted in Section 4.1.3.

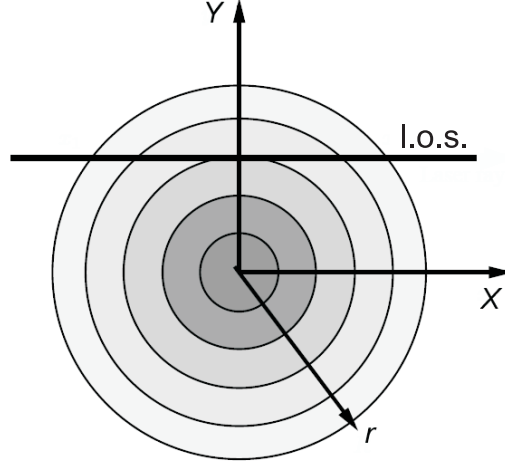


Figure 6.44: Schematic representation of a simple circularly-symmetric magnetic configuration (the Shafranov shift has been neglected), traversed by a horizontal line of sight.

6.5.1.2 Spline representation

A common way of decreasing the number of parameters to be estimated, while imposing a certain adjustable degree of smoothness on the profile, is by expressing the profile in terms of a set of basis functions. We have chosen a set of cubic splines as a basis. Naturally, more sophisticated basis functions can be proposed, and additional boundary conditions can be imposed, but we will not consider this in our first approach to the problem. The spline basis functions pass, for each profile $\eta(r)$, through a set of n_k control points $(\hat{r}_k, \hat{\eta}_k)$. Fixing \hat{r}_k and r_j while varying only $\hat{\eta}_k$, we get η_j by a simple matrix multiplication:

$$\eta_j = \sum_k S_{jk} \hat{\eta}_k.$$

In order to obtain the spline matrix S we calculate the cubic spline $S_k(r)$ passing through (\hat{r}_l, δ_{kl}) (δ_{kl} is the Kronecker delta distribution) and set $S_{jk} = S_k(r_j)$. The line-integrals $L_{\eta,i}$ can then also be calculated through

$$L_{\eta,i} = \sum_{j,k} \widetilde{M}_{ij} S_{jk} \hat{\eta}_k \equiv \sum_{k=1}^{n_k} M_{ik} \hat{\eta}_k,$$

where we have defined the matrix $M = \widehat{M} \times S$. When we speak of the ‘geometry matrix’, we will from now on refer to M .

6.5.1.3 Joint posterior density

We now propose a likelihood function and a prior for the parameters of interest. We start from a set of n_{ff} bremsstrahlung line integrals $L_{\text{ff},i}$, a set of n_{n_e} line-integrated electron densities $L_{n_e,i}$ from interferometry, and a set of n_r CX measurements δ_j , where each δ_j is defined by

$$\delta_j = \sum_k Z_k (Z_k - 1) n_{k,j}.$$

Here, $n_{k,j}$ is the density of impurity species k at the radial position r_j , given in (6.14). In addition, measurements of the local electron temperature $T_{e,j}$ are available, which we assume to be exact. Through the spline matrix S we transform $T_{e,j}$ to $\widehat{T}_{e,k}$. The parameters of interest are the n_k values $\widehat{n}_{e,k}$ and the n_k $\widehat{Z}_{\text{eff},k}$ at the knot positions \widehat{r}_k . We denote the collection of all $\widehat{n}_{e,k}$ (i.e. the complete profile) simply by \widehat{n}_e , and likewise for \widehat{Z}_{eff} , L_{ff} , δ , $L_{n_e,i}$ and \widehat{T}_e . In addition, we will immediately include the possibility of a common scale factor s_{ff} for all line-integrals $L_{\text{ff},i}$, and a common scale factor s_δ for all δ_j . It is possible to consider scale factors that are different for different channels, but we will not do this here. In view of the number of parameters to be estimated, we are working in a $2n_k + 2$ -dimensional probability space. The posterior probability density becomes

$$\begin{aligned} p(\widehat{n}_e, \widehat{Z}_{\text{eff}}, s_{\text{ff}}, s_{n_e} | L_{\text{ff}}, \delta, L_{n_e,i}, \widehat{T}_e, I) \\ \sim \prod_{i_1=1}^{n_{\text{ff}}} \exp \left(-\frac{\gamma_{1,i_1}}{2} \left[L_{\text{ff},i_1} - s_{\text{ff}} \sum_{k_1=1}^{n_k} M_{i_1 k_1} C \bar{g}_{\text{ff}}(\widehat{T}_{e,k_1}) \frac{\widehat{n}_{e,k_1}^2 \widehat{Z}_{\text{eff},k_1}}{\widehat{T}_{e,k_1}^{1/2}} \right]^2 \right) \\ \times \prod_{j=1}^{n_r} \exp \left(-\frac{\gamma_{2,j}}{2} \left[\delta_j - s_\delta \sum_{k_2=1}^{n_k} S_{j k_2} \widehat{n}_{e,k_2} (\widehat{Z}_{\text{eff},k_2} - 1) \right]^2 \right) \\ \times \prod_{i_2=1}^{n_{n_e}} \exp \left[-\frac{\gamma_{\text{I},i_2}}{2} \left(L_{n_e,i_2} - \sum_{k_3=1}^{n_k} N_{i_2 k_3} \widehat{n}_{e,k_3} \right)^2 \right] U_{\widehat{n}_e} U_{\widehat{Z}_{\text{eff}}} U_{s_\epsilon} U_{s_\delta}. \end{aligned}$$

Here, M denotes the geometry matrix for the bremsstrahlung diagnostic, while N is the geometry matrix for the laser beams for interferometry. We have also used the relation (3.3) between bremsstrahlung

emissivity and Z_{eff} . Furthermore, $U_{\hat{n}_e}$ and $U_{\hat{Z}_{\text{eff}}}$ are n_k -dimensional uniform distributions, serving as a prior distribution for the parameters $\hat{n}_{e,k}$ and $\hat{Z}_{\text{eff},k}$, respectively. From the estimated \hat{n}_e and \hat{Z}_{eff} , n_e and Z_{eff} profiles can be calculated using the spline matrix.

Finally, it is possible to model uncertainty on the geometry matrices. To this end, one has to propose estimates for the errors Δ_{ij} on the respective values of the geometry matrix A_{ij} . Then, the variances σ_i^2 on the line-integrated data d_j should be replaced by the effective variances $\sigma_{\text{eff},i}^2$ defined by [113]

$$\sigma_{\text{eff},i}^2 = \sigma_i^2 + \sum_j \Delta_{ij}^2 x_j^2,$$

where x_j represents a component of a general profile. The uncertainty on the geometry matrix is largely determined by uncertainty on the magnetic equilibrium. We have not estimated this uncertainty here, but it should be done in a situation where real physical data are treated.

6.5.1.4 Metropolis sampling

Again an MCMC code was written for the estimation of a \hat{n}_e and \hat{Z}_{eff} profile. An artificial data set was created, based on a typical TEXTOR n_e and T_e profile, and an artificial hollow Z_{eff} profile. The artificial profiles are shown in Figures 6.45, 6.46 and 6.47. Results for differently shaped Z_{eff} profiles are similar. Nine bremsstrahlung line-integrals and six density line-integrals were calculated using the respective geometry matrices. A profile for δ , defined as above, was calculated, and used also as input data. The scale factors s_{ff} and s_{δ} were set to 1.3 and 0.4, respectively. Gaussian noise, described by the respective precisions γ_1 , γ_2 and γ_I , was added to the line-integrated or local measurements. However, in order to retain realistic (line-integrated) profiles, only a few percent noise was included. We used $n_r = 46$, which is usually larger than the number of channels of a CX diagnostic, but the δ profile from CX can be interpolated. The number of knots was chosen to be eight, so we work in a $2 \times 8 + 2 = 18$ -dimensional parameter space. The corresponding basis splines $S_k(r)$ are plotted in Figure 6.48.

The results of the estimation are shown in Figure 6.49 for n_e and Figure 6.50 for Z_{eff} . The MCMC estimated profiles are very close to the original profiles, and the scale factors are also well estimated: $s_{\text{ff}} = 1.31$ and $s_{\delta} = 0.40$. Some oscillation is noticeable in the estimated profiles,

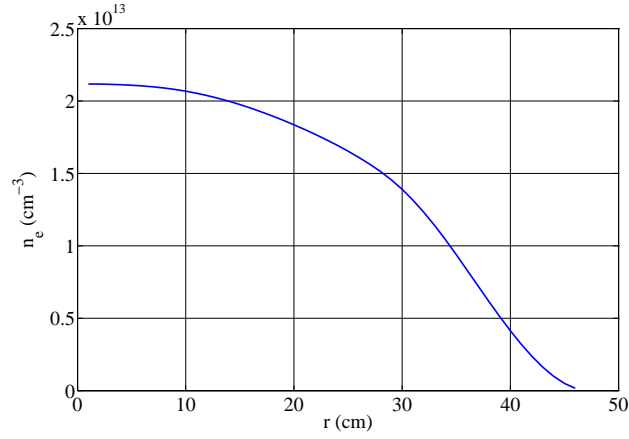


Figure 6.45: An artificial electron density profile used for the MCMC estimation of n_e and Z_{eff} profiles.

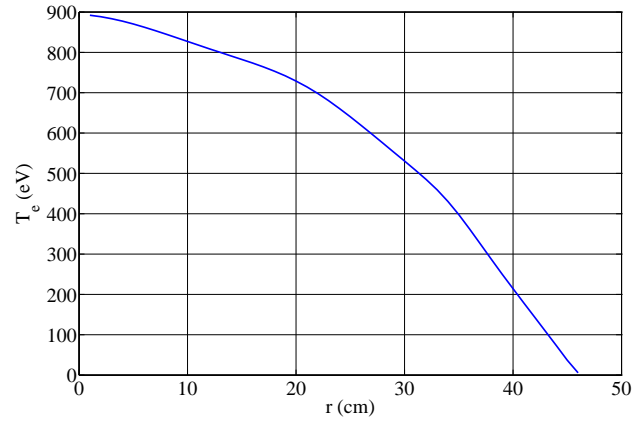


Figure 6.46: An artificial electron temperature profile used for the MCMC estimation of n_e and Z_{eff} profiles.

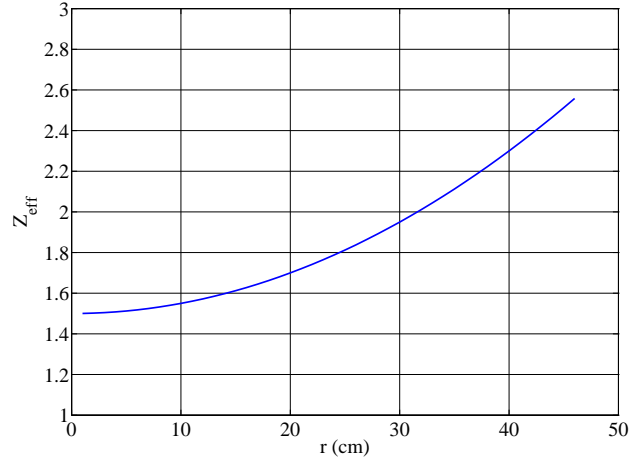


Figure 6.47: An artificial Z_{eff} profile used for the MCMC estimation of n_e and Z_{eff} profiles.

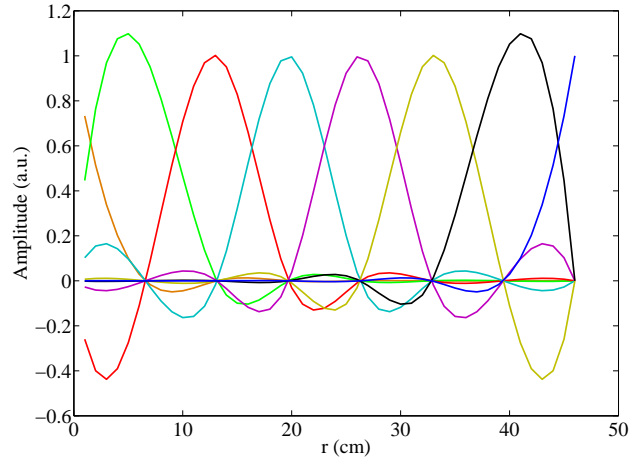


Figure 6.48: The eight spline basis functions that were used in the MCMC estimation of n_e and Z_{eff} profiles.

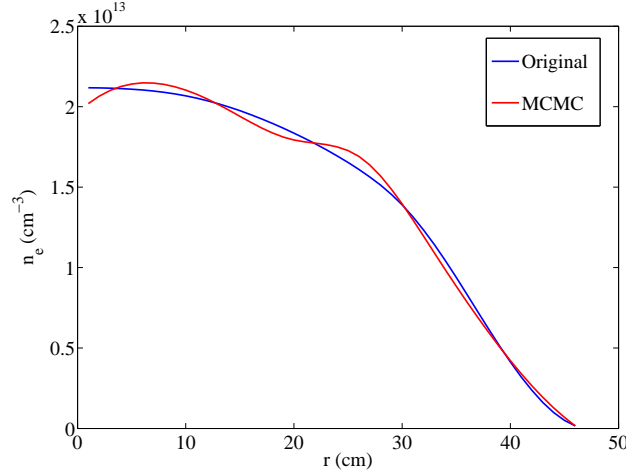


Figure 6.49: MCMC electron density profile estimate from an artificial data set consisting of line-integrated bremsstrahlung emissivity, local CX impurity density and line-integrated electron density. The profile was approximated by a spline.

which is a result of the rather simple set of basis functions used in our approach. As mentioned before, in a later stage a more sophisticated basis can be chosen.

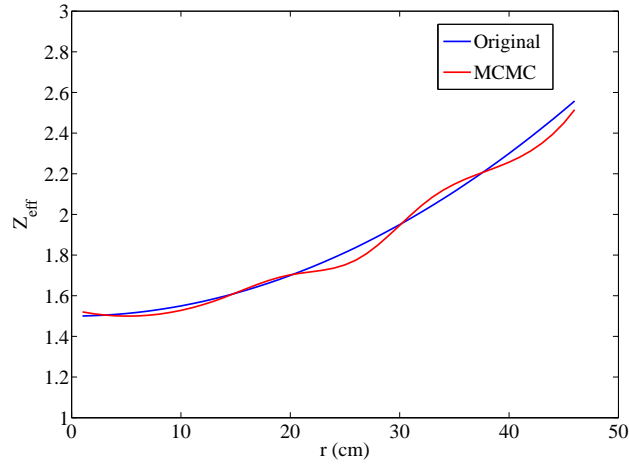


Figure 6.50: MCMC Z_{eff} estimate from an artificial data set consisting of line-integrated bremsstrahlung emissivity, local CX impurity density and line-integrated electron density. The profile was approximated by a spline.

The Feynman Problem-Solving Algorithm:

1. *write down the problem;*
2. *think very hard;*
3. *write down the answer.*

Murray Gell-Mann

7

General conclusions and outlook

In this thesis, we have described our work on the determination of the ion effective charge Z_{eff} in a tokamak plasma. A first part of this work concerns the design and commissioning of a new diagnostic on the TEXTOR tokamak for the derivation of Z_{eff} from bremsstrahlung emissivity measurements in the visible. In a second part, a first step was put towards a Bayesian Integrated Analysis of both measurements of bremsstrahlung emissivity and impurity density from Charge Exchange Spectroscopy. The goal here is to estimate a value for the local Z_{eff} in the plasma, which is consistent with both data sets. We will now present the general conclusions on both subjects. Afterwards, we give

an outlook towards possible improvements to and continuation of the current work.

7.1 Conclusions

7.1.1 Visible bremsstrahlung diagnostic

The new diagnostic on TEXTOR for visible bremsstrahlung emissivity measurement and the derivation of Z_{eff} , was developed entirely within the frame of the present work. The main system is equipped with a CCD camera, while there is also a back-up system served by photomultiplier tubes. The primary merit of the camera system is that a relatively high time resolution can be achieved, while at the same time it features a large number of viewing chords (24), allowing a good reconstruction of emissivity and Z_{eff} profiles. This is a considerable step forward compared to the previous diagnostic.

In addition, and this is also an improvement with respect to the former situation, a line-averaged Z_{eff} value along a centrally viewing chord is now routinely available on-line.

Traditionally, emissivity profiles were reconstructed through Abel inversion of line-integrated signals. Some work has been carried out towards the improved reconstruction of the profiles, particularly as far as the overestimation of Z_{eff} near the plasma boundary is concerned, due to non-bremsstrahlung contributions to the continuum. This is done by regularizing the profiles via Tikhonov and Maximum Entropy regularization.

Finally, in order to improve the calibration that was carried out using an integrating sphere, a novel technique was devised for the calculation of the relative (channel-to-channel) calibration of the diagnostic. Here, the plasma is used as a reference light source. The method can be applied also to other multichannel spectroscopic diagnostics, and has several advantages. The main plus-points are that no dedicated calibration light source is needed, while the calibration can be carried out very easily, without requiring access to the tokamak vessel or the development of a dedicated calibration set-up.

7.1.2 Integrated estimation of Z_{eff}

A number of Bayesian probabilistic models were developed for the estimation of a local Z_{eff} value consistent with both the bremsstrahlung

emissivity and CX carbon density data sets. The first aim was to find a consistent on-axis Z_{eff} value. To this end, time traces for the on-axis measured quantities on JET were used. First only the statistical uncertainties in the input data were modelled, allowing the estimation of Z_{eff} values that are consistent with both input data sets. In principle, explicit electron density data are not needed for the estimation of Z_{eff} , but the introduction of some density measurement increases the precision of the estimated Z_{eff} . However, it was also noticed that, in cases where there is a strong systematic discrepancy between the continuum and CX Z_{eff} , this model appears to be inadequate.

Therefore, in a next step also systematic uncertainties were modelled, by including scale factors for the ϵ and δ data. The parameters of interest were estimated using a Metropolis algorithm, and the method works very well on artificial data. Next, real data from JET discharges were used. It proved to be possible to assess scale factors by which the ϵ and δ data should be multiplied to make the data more consistent. The scaled ϵ and δ are much more mutually consistent than the unscaled versions, to the extent that one might already believe that the correspondence of the estimated Z_{eff} with the real Z_{eff} is rather good. Indeed, we now dispose of two entirely different diagnostic techniques whose measurements are mutually consistent. This also indicates that the proposed model incorporates a relatively good description of the uncertainties. In any case, the error bars are reduced considerably compared to the errors on the continuum and CX Z_{eff} .

In several discharges that were studied (#59193, #59186, #59194), the estimated Z_{eff} is more similar, both qualitatively and quantitatively, to the CX Z_{eff} than to the continuum Z_{eff} . The ϵ data (hence also ϵ_{ff}), and thus the continuum Z_{eff} , appear to be consistently overestimated by an average factor of 1.4. This is perhaps a bit surprising, since a general feeling among diagnosticians is that, on the contrary, the continuum Z_{eff} usually is the most accurate of the two. Our analysis raises some doubt regarding this general belief. The questions rise as to what the reasons are for this preference of the MCMC estimated Z_{eff} for the CX Z_{eff} , and whether, indeed, it is a general feature of the estimation results. The former may be answered by increasing the model sophistication (see the outlook section), whereas the latter question suggests a similar systematic analysis of many discharges. A possible explanation of a consistent bremsstrahlung emissivity overestimation might be found in the influence of non-bremsstrahlung contributions to the con-

tinuum, or perhaps in residual in-vessel reflections or in the occurrence of atomic lines in the observed spectral window.

In some of the considered discharges (#61352, #59193, #59186), the δ data, i.e. the summed impurity density, appear to be (seriously) underestimated (factor 0.4 to 0.8). In the case where only fully stripped carbon makes an appreciable contribution to Z_{eff} , this yields an estimate for the degree of underestimation of the carbon density. If also other impurities make a non-negligible contribution to Z_{eff} , then one can still hold on to the proposed model if the densities of the other impurities are simply related to the carbon density by a factor. One can then, for example, understand the underestimation of δ in terms of the influence of not monitored impurities.

Finally, a Bayesian model was proposed that permits the estimation of Z_{eff} profiles, consistent with measurements of line-integrated bremsstrahlung emissivity, local impurity density from CXS and line-integrated electron density. This method requires knowledge of the geometry matrix associated with the lines of sight of the respective diagnostics for bremsstrahlung and electron density measurement. The profiles of interest were expressed in terms of a spline basis. A common scale factor was introduced for all channels of the bremsstrahlung diagnostic, and another common scale factor for the channels of the CX diagnostic. The estimation works well on artificial data.

In the present context, a question that is often heard is why one should attach more value to the Bayesian estimate of Z_{eff} as compared to the continuum or CX Z_{eff} . The answer lies in the observation that both the derived continuum and CX Z_{eff} are subject to a lot of uncertainty. In such a case, the best one can do is to take into account all known data and uncertainties, and calculate a ‘most probable’ (or average) Z_{eff} on that basis. If the uncertainties are modelled to a sufficiently detailed extent, then it is indeed expected that the estimated Z_{eff} is closer to the real physical Z_{eff} , as far as this information is contained in the data and the physical model. And of course, as mentioned before, the fact that the input data can be made more consistent by modelling the uncertainties, is a good indication that the ‘truth’ has been approached better.

7.2 Outlook

We here present an overview of the main points in the work described in this thesis that can be improved, or which suggest further directions of the research.

7.2.1 Visible bremsstrahlung diagnostic

A few minor issues can be identified in the design of the visible bremsstrahlung diagnostic on TEXTOR. For example, it could prove to be valuable if the field of view were widened. However, the main future interests lie somewhere else.

To begin with, a further study is necessary of the performance of Tikhonov and Maximum Entropy regularized inversion of the bremsstrahlung emissivity line-integrals for the reconstruction of emissivity profiles. Specifically, it would be very useful to obtain more information on the propagation characteristics and magnitude of systematic uncertainties at the plasma edge. The Maximum Entropy method allows to impose prior knowledge about the emissivity profile, which could possibly still improve the properties of the inversion.

Next, the possibilities of the relative calibration of the system using requirements of profile consistency, should be explored in more detail. Additional test discharges are required, estimating also error bars on the calibration factors. The alternative method, where not the plasma is shifted, but rather the fibre support is rotated, should be tried. Finally, it would be very interesting to apply this method also to the calibration of the upgraded diagnostic for carbon spectroscopy (CIII and CV) on TEXTOR. This would replace a difficult to perform calibration, where a dedicated UV light source is needed, by a very easy and convenient procedure. As mentioned before, this will be attempted in autumn 2006.

7.2.2 Integrated estimation of Z_{eff}

The integrated estimation of Z_{eff} , as described in this thesis, has only started, and many improvements to the technique can be suggested.

First, for the method that takes into account both statistical and systematic uncertainties for the estimation of the on-axis Z_{eff} , we would like to set up a database of test results in JET discharges. This would permit to make statements about the quality of the continuum and

CX Z_{eff} estimates in different discharge scenarios, while average values for both the s_ϵ and s_δ scale factors could be assessed. Here, it would be useful, where possible, to include CX measurements of the densities of other light impurities, not only carbon. On the other hand, more information on the edge continuum uncertainty would be advantageous also here. In addition, the method should also be applied to the off-axis regions of the plasma cross-section.

Next, the method for the estimation of Z_{eff} profiles should be tested on a real data set. Afterwards also here a database of estimation results should be established.

It would also be valuable to conduct the analyses on TEXTOR data, in order to study the effect of the DED on Z_{eff} . This will be attempted in autumn 2006.

In the longer run, the sophistication of both models should be increased. The eventual aim of this research is to be able to estimate full Z_{eff} profiles, consistent with both the bremsstrahlung and CX data, starting from the raw, uncalibrated measurements of the respective detectors. To this end, all uncertainties of importance should be modelled: statistical uncertainties in proper likelihood PDFs, systematic uncertainties as nuisance parameters. Prior knowledge on Z_{eff} may be refined, depending on the plasma conditions, possibly also based on other experiments. This will result in a very complicated posterior distribution, with many nuisance parameters that all have to be marginalized. The use of MCMC sampling methods will therefore be essential. A sensitivity analysis will then allow to assess the impact of the main sources of uncertainty in the data descriptive model on the Z_{eff} error bars. As a result, the most critical sources of uncertainty can be properly addressed. In addition, Bayesian model selection can help to find the model that sufficiently describes the data, while minimizing the number of required parameters.

On the other hand, the techniques presented in the present work can be applied to the estimation of other plasma quantities as well. In particular, a Bayesian model for the estimation of impurity concentrations from CXS would be very useful in itself. One of the primary applications should be the improvement of the reliability and robustness of CX measurements of helium concentrations, which is critical for helium exhaust studies for fusion reactors.

Not only does Bayesian analysis of fusion diagnostic data allow to improve to a considerable extent the quality of estimates for physical

parameters, but also does it become possible to optimize the design of the diagnostics involved in the measurements, *in function of the parameters of interest*. To this end, sensitivity analyses combined with the concept of Bayesian Diagnostic Design should be applied to CXS. In particular, this can assist in the optimization of the new diagnostic for active beam spectroscopy on TEXTOR, which is a pilot experiment for the ITER active beam diagnostic.

Bibliography

- [1] F. Joos, *Europhysics News* **27**, 213 (1996).
- [2] P. Stangeby, *The plasma boundary of magnetic fusion devices*, Institute of Physics Publishing, Bristol, 2000.
- [3] M. von Hellermann and H. Summers, *Active Beam Spectroscopy at JET*, pages 135–164, Elsevier Science Publishers, London, 1993.
- [4] J. Ongena and G. Van Oost, *Transactions of Fusion Technology* **37**, 3 (2000).
- [5] G. Van Oost, G. Verdoolaege, and J. Ongena, *Tijdschrift VeLeWe (Vereniging Leraars Wetenschappen)* (2002).
- [6] U. N. P. Division, *World population prospects: the 2004 revision*, Technical report, United Nations, New York, 2004.
- [7] A. Camplani and A. Zambelli, *Endeavour, New Series* **10**, 132 (1986).
- [8] M. Golay and N. Todreas, *Sci. Am.* **262**, 58 (1990).
- [9] R. Gelbspan, *The heat is on: the climate crisis, the cover-up, the prescription*, Perseus Books, Reading Massachusetts, updated edition, 1998.
- [10] BP statistical review of world energy, 2006, <http://www.bp.com/>.
- [11] J. Raeder et al., *Safety and environmental assesment of fusion power (SEAFP)*, Technical Report EURFUBRU XIII-217/95, European Commission, Brussels, Belgium, 1995.

-
- [12] S. Barabaschi et al., Fusion programme evaluation 1996, Technical Report EUR 17521, European Commission, Brussels, Belgium, 1996.
 - [13] J. Wesson, *Tokamaks*, Oxford University Press, New York, second edition, 1997.
 - [14] B. Kadomtsev, *Tokamak plasma: a complex physical system*, IOP Publishing, London, 1992.
 - [15] R. Dendy, *Plasma physics, an introductory course*, Cambridge University Press, Cambridge, 1995.
 - [16] J. Sheffield, Rev. Mod. Phys. **66**, 1015 (1994).
 - [17] G. McCracken and P. Stott, Nucl. Fusion **191**, 889 (1979).
 - [18] R. Isler, Nucl. Fusion **24**, 1599 (1984).
 - [19] R. Jensen, D. Post, W. Grasberger, C. Tarter, and W. Lokke, Nucl. Fusion **17**, 1187 (1977).
 - [20] A. Messiaen et al., Phys. Plasmas **4**, 1690 (1997).
 - [21] J. Winter, J. Nucl. Mater. **176-177**, 14 (1990).
 - [22] J. Winter, Plasma Phys. Control. Fusion **38**, 1503 (1996).
 - [23] S. Hirschman and D. Sigmar, Nucl. Fusion **21**, 1079 (1981).
 - [24] R. Balescu, *Transport Processes in Plasmas: Classical Transport Theory*, Elsevier Science, London, 1988.
 - [25] R. Balescu, *Transport Processes in Plasmas: Neoclassical Transport Theory*, Elsevier Science, London, 1988.
 - [26] R. Balescu, *Aspects of Anomalous Transport in Plasmas*, Taylor & Francis, London, 2005.
 - [27] D. Reiter, G. Wolf, and H. Kever, Nucl. Fusion **30**, 2141 (1990).
 - [28] R. Behrisch and V. Prozesky, Nucl. Fusion **30**, 2166 (1990).
 - [29] I. Hutchinson, *Principles of Plasma Diagnostics*, Cambridge University Press, Cambridge, second edition, 2002.

-
- [30] R. Goldston and P. Rutherford, *Introduction to Plasma Physics*, IOP Publishing Ltd., London, 1995.
 - [31] S. Seshadri, *Fundamentals of Plasma Physics*, American Elsevier Publishing Company, New York, 1973.
 - [32] R. Goldston and P. Rutherford, *Fluid equations for a plasma*, chapter 6, In [30], 1995.
 - [33] J. Wesson, *Plasma physics*, chapter 2, In [13], second edition, 1997.
 - [34] L. Spitzer, Jr. and R. Härm, Phys. Rev. **89**, 977 (1953).
 - [35] R. Goldston and P. Rutherford, *Collisions in fully ionized plasmas*, chapter 11, In [30], 1995.
 - [36] J. Wesson, *Equilibrium*, chapter 3, In [13], second edition, 1997.
 - [37] J. Wesson, *Appendix*, chapter 14, In [13], second edition, 1997.
 - [38] J. Wesson, *Heating*, chapter 5, In [13], second edition, 1997.
 - [39] D. Maisonnier et al., A conceptual study of commercial fusion power plants: final report of the European Fusion Power Plant Conceptual Study (PPCS), Technical Report EFDA-RP-RE-5.0, EFDA, Brussels, Belgium, 1996.
 - [40] G. Telesca, H. Claassen, A. Pospieszczyk, R. Koch, and D. Van Eester, Nucl. Fusion **34**, 625 (1994).
 - [41] A. Boileau, M. von Hellermann, L. Horton, J. Spence, and H. Summers, Plasma Phys. Control. Fusion **31**, 779 (1989).
 - [42] K. Kadota, M. Otsuka, and J. Fujita, Nuclear Fusion **20**, 209 (1980).
 - [43] T. Boyd and J. Sanderson, *Plasma radiation*, chapter 9, Thomas Nelson and Sons, London, 1969.
 - [44] I. Hutchinson, *Electromagnetic emission by free electrons*, chapter 5, In [29], second edition, 2002.
 - [45] C. Wyatt, *Radiometric Calibration: Theory and Methods*, Academic Press, New York, 1978.

-
- [46] A. Sommerfeld, *Atombau und Spektrallinien*, Friedrich Vieweg und Sohn, Braunschweig, 1939.
 - [47] P. Brussaard and H. Van de Hulst, *Rev. Mod. Phys.* **34**, 507 (1962).
 - [48] H. Griem, *Plasma Spectroscopy*, chapter 5, McGraw-Hill, Inc., New York, 1964.
 - [49] A. Ramsey, *Rev. Sci. Instrum.* **58**, 1211 (1987).
 - [50] W. Karzas and R. Latter, *Astroph. J. Suppl.* **6**, 167 (1961).
 - [51] N. Schoon et al., Measurement of the bremsstrahlung radiation in the visible and determination of Z_{eff} on Torus Experiment for Technology Oriented Research (TEXTOR), Technical Report 92, LPP – Royal Military Academy, Brussels, Belgium, 1990.
 - [52] M. von Hellermann et al., *Plasma Phys. Control. Fusion* **37**, 71 (1995).
 - [53] M. von Hellermann et al., *Review of Scientific Instruments* **61**, 3479 (1990).
 - [54] N. Hawkes, *Experimental studies of ion pressure, impurity flows and their influence on transport in the JET tokamak*, PhD thesis, Imperial College London, 1995.
 - [55] M. von Hellermann, A charge exchange analysis package for ITER, in *Proceedings of the 3rd Workshop on Fusion Data Processing, Validation and Analysis*, Cadarache, 2004.
 - [56] A. Costley et al., ‘Review on diagnostics for magnetic fusion’, to be published in *Fusion Science and Technology* (2006).
 - [57] M. von Hellermann, private communication.
 - [58] B. Schunke, G. Huysmans, and P. Thomas, *Rev. Sci. Instrum.* **76**, 083501 (2005).
 - [59] G. Verdoolaege, G. Telesca, E. Delabie, G. Van Oost, and the TEXTOR team, *Rev. Sci. Instrum.* **77**, 10F310 (2006).
 - [60] A. Donné et al., *Fusion Science and Technology* **47**, 220 (2005).

- [61] G. Holst, *CCD Arrays, Cameras, and Displays*, SPIE – International Society for Optical Engineering, Bellingham, WA, 1998.
- [62] J. Krom, The Trilateral Euregio Cluster (TEC) Web-Umbrella, transparent access to experimental and derived data, in *Proceedings of the 4th Workshop on Fusion Data Processing, Validation and Analysis*, Jülich, 2006.
- [63] G. Telesca, G. Verdoolaege, K. Crombé, B. Unterberg, and G. Van Oost, Screening and radiation efficiency of carbon with Dynamic Ergodic Divertor on TEXTOR, in *Proceedings of the 32nd EPS Conference on Plasma Physics and Controlled Fusion*, P-1.020, Tarragona, 2005.
- [64] G. Telesca et al., J. Nucl. Mater. **337-339**, 361 (2005).
- [65] G. Telesca et al., Investigation of carbon screening on TEXTOR with Dynamic Ergodic Divertor in 6/2 mode, in *Proceedings of the 33rd EPS Conference on Plasma Physics and Controlled Fusion*, P-2.157, Rome, 2006.
- [66] G. Verdoolaege, G. Telesca, G. V. Oost, and the TEXTOR team, Optimization of the relative calibration for a visible bremsstrahlung Z_{eff} diagnostic on TEXTOR via requirements of profile consistency, in *Proceedings of the 33rd EPS Conference on Plasma Physics and Controlled Fusion*, P-2.165, Rome, 2006.
- [67] G. Telesca, G. Van Oost, J. Ongena, C. Leys, and N. Schoon, Investigation of the sensitivity of visible continuum $Z_{\text{eff}}(0)$ to edge radiation, in *Proceedings of the ICPP 1992*, volume II, page 1147, Innsbruck, 1992.
- [68] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, Springer-Verlag, New York, 2004.
- [69] P. Hansen and D. O’Leary, SIAM J. Sci. Comput. **14**, 1487 (1993).
- [70] G. Cottrell, *Maximum entropy and plasma physics*, chapter 5, Oxford University Press, New York, 1991.
- [71] M. E. Foord, E. S. Marmar, and J. L. Terry, Rev. Sci. Instrum. **53**, 1407 (1982).

-
- [72] D. Kalupin et al., Plasma Physics and Controlled Fusion **45**, 1501 (2003).
 - [73] D. Kalupin et al., Dependence of confinement on fuel isotope in impurity seeded plasmas, in *Proceedings of the 30th EPS Conference on Plasma Physics and Controlled Fusion*, St. Petersburg, P-2.136, 2003.
 - [74] B. Unterberg et al., Overview of transport studies with the Dynamic Ergodic Divertor in the tokamak TEXTOR, in *Proceedings of the 2nd Workshop on Stochasticity in Fusion Plasmas*, Jülich, 2005.
 - [75] A. Boileau, M. von Hellermann, L. Horton, H. Summers, and P. Morgan, Nuclear Fusion **29**, 1449 (1989).
 - [76] H. Weisen et al., Nuclear Fusion **29**, 2187 (1989).
 - [77] R. Cox, American Journal of Physics **14**, 1 (1946).
 - [78] C. Caves, Why we should think of quantum probabilities as Bayesian probabilities, in *Proceedings of the 26th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Paris, 2006.
 - [79] A. Gelman, J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*, Chapman & Hall/CRC, Boca Raton, second edition, 2004.
 - [80] D. Sivia, *Data Analysis: a Bayesian Tutorial*, Oxford University Press, New York, 1996.
 - [81] P. Gregory, *Bayesian logical data analysis for the physical sciences*, Cambridge University Press, Cambridge, 2005.
 - [82] D. MacKay, Neural Computation **4**, 415 (1992).
 - [83] C. Rodríguez, Data, virtual data and anti-data, in *Proceedings of the 26th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Paris, 2006.
 - [84] C. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, New York, second edition, 2004.

- [85] W. Gilks, S. Richardson, and D. Spiegelhalter, *Monte Carlo Statistical Methods*, Chapman & Hall, London, 1996.
- [86] D. Gamerman, *Markov Chain Monte Carlo: stochastic simulation for Bayesian inference*, Chapman & Hall, London, 1997.
- [87] N. Metropolis and S. Ulam, *Journal of the American Statistical Association* **44**, 335 (1949).
- [88] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, *Journal of Chemical Physics* **21**, 1087 (1953).
- [89] W. Hastings, *Biometrika* **57**, 97 (1970).
- [90] A. Gelfand and A. Smith, *Journal of the American Statistical Association* **85**, 398 (1990).
- [91] S. Geman and D. Geman, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721 (1984).
- [92] A. Gelman, *Computing science and statistics* **24**, 433 (1992).
- [93] L. Tierney, *The Annals of Statistics* **22**, 1701 (1994).
- [94] S. Meyn and R. Tweedie, *Markov chains and stochastic stability*, Springer-Verlag, New York, 1993.
- [95] G. Verdoolaege, M. von Hellermann, R. Jaspers, M. Ichir, and G. Van Oost, Integrated bayesian estimation of Z_{eff} in the TEXTOR tokamak from bremsstrahlung and CX impurity density measurements, in *Proceedings of the 26th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Paris, 2006.
- [96] G. Verdoolaege, G. Telesca, G. Van Oost, and the TEXTOR team, Integrated Bayesian estimation of Z_{eff} on TEXTOR from bremsstrahlung and CX impurity density measurements, in *Proceedings of the 4th Workshop on Fusion Data Processing, Validation and Analysis*, Jülich, 2006.
- [97] G. Verdoolaege, G. Telesca, and G. Van Oost, Bayesian estimation of Z_{eff} in the TEXTOR tokamak from visible bremsstrahlung and Charge Exchange Recombination Spectroscopy (CXRS) measurements, in *Proceedings of the 25th International Workshop on*

- Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, page 456, San José CA, 2005.
- [98] G. Verdoolaege, G. Telesca, and G. Van Oost, Reconstruction of Z_{eff} profiles at TEXTOR through Bayesian source separation, in *Proceedings of the 12th International Congress on Plasma Physics*, Nice, CCSD-00001802, 2004.
- [99] R. Fischer et al., Plasma Phys. Control. Fusion **44**, 1501 (2002).
- [100] H. Dreier et al., ‘Design of diagnostics: Case studies for Wendelstein 7-X’, to be published in Fusion Science and Technology (2006).
- [101] R. Fischer, A. Dinklage, and E. Pasch, Plasma Phys. Control. Fusion **45**, 1095 (2003).
- [102] R. Fischer and A. Dinklage, Rev. Sci. Instrum. **75**, 4237 (2004).
- [103] A. Dinklage, R. Fischer, and J. Svensson, Fusion Science and Technology **46**, 355 (2004).
- [104] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, 2001.
- [105] G. Verdoolaege, G. Telesca, and G. Van Oost, Reconstruction of Z_{eff} profiles at the TEXTOR tokamak through Bayesian source separation, in *Proceedings of the 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, page 344, Garching bei München, 2004.
- [106] G. Verdoolaege, G. Telesca, G. Van Oost, and G. Van Den Berge, Reconstruction of the edge Z_{eff} profile from bremsstrahlung data via extensions of Independent Component Analysis on TEXTOR, in *Proceedings of the 29th EPS Conference on Plasma Physics and Controlled Fusion*, Montreux, P-2.060, 2002.
- [107] G. Verdoolaege, G. Telesca, and G. Van Oost, Neural network methods for radial profile reconstruction for Z_{eff} from bremsstrahlung data on the TEXTOR tokamak, in *Proceedings of the 5th International FLINS Conference on Computational Intelligent Systems for Applied Research*, page 615, Gent, 2002.

-
- [108] G. Verdoolaege, G. Telesca, and G. Van Oost, Reconstruction of edge Z_{eff} profiles from bremsstrahlung data via (semi-)Blind Source Separation methods on TEXTOR, in *Verhandlungen der Deutschen Physikalischen Gesellschaft*, Aachen, P 10.21, 2003.
 - [109] G. Matthews et al., *J. Nucl. Mater.* **241–243**, 450 (1997).
 - [110] G. Telesca et al., *Nucl. Fusion* **36**, 347 (1996).
 - [111] G. Roberts, A. Gelman, and W. Gilks, *Anal. of Applied Probability* **7**, 110 (1997).
 - [112] R. Neal, Probabilistic inference using Markov Chain Monte Carlo methods, 1993, <http://omega.albany.edu:8008/neal.pdf>.
 - [113] V. Dose, R. Fischer, and W. von der Linden, *Deconvolution based on experimentally determined apparatus functions*, page 147, Kluwer Academic, Dordrecht, 1998.