

Master Thesis: Optimization of Hybrid
Scenario ITER Ramp-Up

Jeroen van Dongen
Student Theoretical Physics
Studentnumber 3275299
Utrecht University

August 5, 2013

Contents

1	Introduction	3
1.1	Tokamak Device	5
1.2	Simulation Codes	6
1.3	Research Questions	6
2	Tokamak Physics	8
2.1	Grad-Shafranov equation	10
2.2	Flux Averaging and Poloidal Flux Quantities	11
2.3	Transport Equations	13
2.3.1	Poloidal Flux Diffusion	13
2.3.2	Particle Transport	15
2.3.3	Energy Transport	16
2.4	Sources	16
2.5	Heating Systems	17
2.6	Fusion Reaction	22
3	Theory of Optimization	26
3.1	Direct Search Method	28
3.2	Line Search Method	29
3.2.1	Step Length	30
3.2.2	Convergence of Line Search Methods	33
3.3	Sequential Quadratic Programming	35
4	Alpha Particle Heating	37
4.1	Direct Deposition Model	38
4.2	Simple Diffusion Model	38
4.3	Benchmark Results	40
5	Validating The RAPTOR Code For ITER	43
5.1	Introduction	43
5.2	Baseline Comparison	45

5.2.1	Transport Model: Bohm-GyroBohm	46
5.2.2	Conductivity	46
5.2.3	Evolution Comparison	48
5.3	Poloidal Flux Diffusion	51
5.3.1	Bootstrap Current	52
5.4	Energy Transport Equation	52
5.4.1	Particle Flux	52
5.4.2	Ion-Electron Heat Exchange	54
5.4.3	Radiation Losses	56
5.4.4	Auxiliary Heating	59
5.4.5	Alpha Particle Heating	65
5.5	Full Transport Model Comparison	66
5.5.1	Electron-Ion Heat Exchange	66
6	Optimization	67
6.1	Optimization set-up	67
6.1.1	Cost-Function	68
6.1.2	Constraints	70
6.1.3	Local Minima	72
6.2	L-mode Optimization Results	73
6.2.1	Reference Case	73
6.2.2	Optimized Trajectories	75
6.2.3	Model Sensitivity	84
6.2.4	Active Constraints	86
6.3	H-mode Optimization	87
7	Conclusion	94
7.1	Outlook	94
7.2	Acknowledgements	95
A	Mathematical Derivations	96
A.1	Flux Diffusion	96
A.2	Particle Diffusion	98
B	Bohm-GyroBohm Transport Model	100

Chapter 1

Introduction

Utilizing nuclear fusion for energy consumption purposes has been a topic of research since the early 1920's when it was discovered that 2 hydrogen atoms were heavier than 1 helium atom. Ever since the realization that the sun burned up hydrogen to produce power that counteracts gravity it has been the goal of many scientists to build a nuclear fusion reactor that could supply mankind with energy. Being able to exploit most abundant energy supply method in the universe would mean a safe and overly abundant energy supply. Using 1 kg of hydrogen plasma fuel would release 10^8 kWh, equivalent to running a 1 GW power station for a day. It has been reported that the earth's natural Lithium abundance would be able to supply us with enough Tritium for several of hundreds of years. Deuterium is a stable isotope and the Earth's oceans are comprised of 0.0153% of Deuterium making it accessible for thousands of year

During all the years of research it has become clear that achieving a profitable fusion reactor would not be easily achieved. Many generations of scientists have been set back by the discovery of plasma instabilities, making the confinement of the plasma more and more problematic. Two main schemes are used to develop a fusion reactor. First is the magnetic confinement reactor and secondly the inertial confinement. This thesis will focus solely on the first method.

The reaction that achieves fusion easiest is with two isotopes of hydrogen: Deuterium and Tritium which have 1 and 2 neutrons respectively. They produce an alpha (helium) particle and a free neutron:



The reaction is induced through collisions between these particles and it is therefore very important to know the cross-section. The cross section

for different hydrogen isotopes can be seen in figure 1.1. It is clear that the cross-section between Tritium and Deuterium is greatest and has a maximum around 100 keV. The required reactions occur in the high energy tail of the Maxwellian distribution and a plasma temperature of around 10 keV (100 million degrees centigrade) is therefore sufficient for fusion reactions. This can be seen in figure 1.2a and the final result for the total reaction rate can be seen in figure 1.2b. A large fusion reaction rate is not the only important

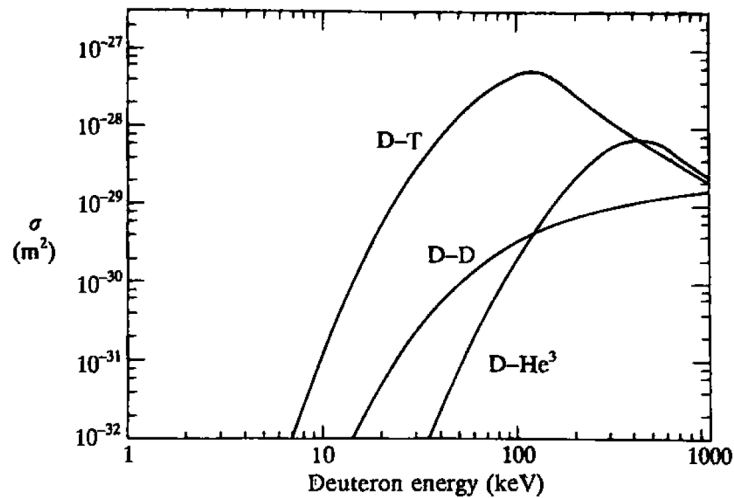
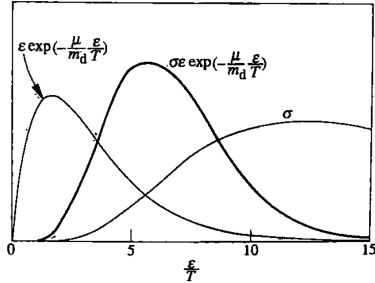


Figure 1.1: *The fusion cross sections for different hydrogen isotopes*

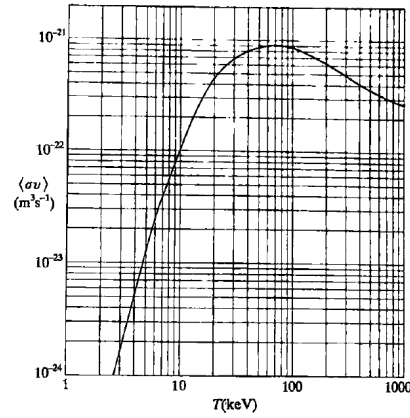
ingredient for net positive energy generation. To ensure the reaction becomes self sustainable the energy released in the fusion reaction must be redeposited into the plasma. This requires that the energy is confined for long enough in a region sufficiently close to the active fusion region. This is expressed to the so called *energy confinement time* (τ_E) defined by the ration between total plasma energy and the power loss. An important goal for fusion plasma physics is ignition. This means that the power generated by fusion reactions is sufficient to overcome the power losses and thus maintain the plasma in the burning regime. This goal is formulated by the well-known Lawson criterion for flat density and temperature profiles (1.1):

$$nT\tau_E \geq 3 \times 10^{21} \text{ keVs}, \quad (1.1)$$

where n is the density and T is the temperature of the plasma in keV. If a typical tokamak temperature of 20 keV is used, we can require that the $n\tau_E > 1.5 \times 10^{20} \text{ m}^{-3}\text{s}$ to reach ignition.



(a) The fusion cross sections for different hydrogen isotopes



(b) The total reaction rate for a Tritium-Deuterium fusion reaction

1.1 Tokamak Device

Within the two main research routes different approaches are being used. In the inertial confinement research a high density is created but the confinement time is poor whereas in the magnetic confinement scheme low density plasma's ($n > 10^{20} \text{m}^{-3}$) are confined for a longer period ($\tau_E > 1.5 \text{s}$). The most generally accepted form of magnetic confinement is the geometry of a tokamak (T² Torus) which was designed by Russian scientists in the 1960's. The magnetic fields in this device are generated (mostly) by the electrical currents in the plasma itself. Unfortunately this plasma current is also responsible for creating instabilities through magnetic energy. This limits the maximally achievable confinement time through global plasma instabilities or locally enhanced transport. These instabilities have to be avoided as much as possible but will not be the topic of this thesis. They will however be briefly considered during the next chapters.

Another important downside to the tokamak geometry comes from the method of generating the plasma current. At this point the plasma current is generated through inductance with a primary coil and using the plasma as a secondary coil. By ramping up the current in this primary coil a current is induced in the plasma, the magnitude being proportional to the current ramp up rate. The downside of the tokamak geometry is that the primary coil current can not be ramped up indefinitely. Therefore either the current driving force has to be replaced at some point of the plasma operation by a noninductive one or the tokamak has to be operated using a pulsed scheme. Usually full noninductive current drive is not achievable but the

time of operation is enlarged by using various noninductive current driving schemes.

1.2 Simulation Codes

1.3 Research Questions

1. **Can the RAPTOR code be adapted to effectively model the physics in ITER and be an effective way to compute optimal plasma scenarios?**

This question arises because RAPTOR was formerly used to model smaller sized tokamaks and at the beginning of the work on this thesis not yet equipped with the proper physics to model an ITER scenario. The question has two important aspects that need to be considered. First is whether the code can be adapted such that it models the appropriate physics to such an extent that it produces results that are representative to the real world physics. The second question is whether this can be done within the computational speed needed to perform heuristic optimization, as opposed to the now more primitive approach that is standard in the fusion physics society. When these two questions can be answered with a positive reply there will be a basis to start looking for new optimized trajectories that might possibly hold more efficient possibilities.

2. **How can alpha particle production and heating be modeled within the RAPTOR code.**

One of the important processes that were not yet included in RAPTOR was the production of alpha particles through fusion reactions. These alpha particles can be a considerable contribution to the total heating of the plasma and also form the main goal of fusion reactors in the future. The question is whether the deposition profile can be effectively modeled in RAPTOR by making use of several assumptions. The problem with this process is the cooling down of alpha particles by collisions with the plasma constituents. This is a process that is complicated by diffusion of the energetic particles on the same time scale as on which they lose their energy to the plasma. This diffusion would imply considerable extra computation to correctly model the process. This would greatly diminish the effectiveness of the RAPTOR code. Therefore assumptions have to be made to decrease computation time and these assumptions need to be tested. The question remains whether these

assumptions to not constraint the model to be inconsistent with the real world physics.

3. Do the optimal trajectories proposed by RAPTOR yield a difference from the previously obtained trajectories through CRONOS?

The main objective of this thesis is to find new trajectories which will more effectively heat an ITER discharge towards a hybrid scenario flat-top. Previous work is based mostly on 'turn the knob' optimization which bases its effectiveness on the expertise of the researcher. RAPTOR provides a way to heuristically optimize the trajectories used for heating a discharge at the cost of the level of detail in the physics. The question is thus whether an updated version of RAPTOR will be able to provide trajectories that yield better results than the ones previously obtained. The problems here could be that the resulting profiles as a result of trajectory optimization is different in CRONOS. Another problem could be that constraints or desirable trajectory traits can not be included in the optimization procedure. Next there is a problem with the cost function sensitivity. When the cost function landscape is too flat in any direction, the optimization problem can taken considerable amount of iterations. This is illustrated in the optimization of the Rosenbrock function which has a long stretched valley, making it difficult to reach the global minimum with standard optimization procedures. The cost function for the problem at hand is a function of many variables and therefore difficult to predict whether such a flat landscape exists. This fact also prevents us from being able to make any conclusion whether the solution is also the global solution. This thesis will also try to find ways to maximize the chance that the solution is a global solution.

Chapter 2

Tokamak Physics

This chapter will deal in greater detail with the physics of plasma's confined in a tokamak geometry. A lot of which will be based on the work in the Tokamak physics book by John Wesson[1]. First some general plasma properties will be discussed after which more specialized physics will be discussed such as tokamak heating and diffusion.

The three traditional states of matter (solids, liquids and gases) are build up of neutral atoms. When the temperature of a gas is increased beyond the ionization limit the electrons and positively charged nuclei are separated. The transition happens in a narrow region of temperature and is very often called a phase transition into the plasma phase. In the plasma state the kinetic energy exceeds the potential energy of the electrostatic coulomb energy. The particles in the plasma typically collide quasi-elastically through their coulomb interactions but when ions become sufficiently close together there is a possibility that they tunnel through the coulomb barrier and fuse to form a heavier nuclei. The quasi-neutrality condition states that $\eta_e \simeq \sum_j Z_j \eta_j$ and still assumes that there are sources of electric fields: $\nabla \cdot \mathbf{E} \neq 0$. Even though the plasma tends to this quasi-neutrality because of coulomb attraction between opposite charge, random thermal motion will cause small charge fluctuations. These fluctuations will be limited by the fact that their potential energy due to their separation can not be much larger then their thermal energy. This leads to a screening on length scales smaller then the characteristic length scale called the Debye length: $\lambda_D \simeq \sqrt{\epsilon_0 k_b T / e^2 n}$. Similarly a characteristic time scale can be calculated called the plasma frequency: $\omega_p = \sqrt{n_e e^2 / \epsilon_0 m_e}$. The ions do not contribute to this time scale because of there large mass. The plasma frequency is temperature independent and electromagnetic waves with frequencies above this frequency can propagate through the plasma. Electromagnetic waves with $\omega < \omega_p$ will be screened by the electrons and will exponentially decay.

The fact that all constituents of the plasma are charged makes them susceptible to the Lorentz force. When the plasma is placed in a magnetic field the particles are free to move parallel to the magnetic field while perpendicular motion is restricted to gyration around the magnetic field lines. Therefore the particles are more or less confined by a magnetic field. The gyration radius is given by $\rho_s = \frac{m_s v_\perp}{e Z_s B}$ and it can be seen that the larger the magnetic field the smaller the gyration radius. Thus this leads to a stronger confinement. A tokamak keeps the plasma in an axis-symmetric toroidal magnetic chamber by using strong helical magnetic fields.

The particles do however drift across the magnetic field when they are subjected to external forces or field inhomogeneities. This is a mechanism that results in confinement losses. The first is the drift velocity due to a perpendicular electric field:

$$\mathbf{v}_E = \frac{\mathbf{E} \times \mathbf{B}}{B^2}, \quad (2.1)$$

which is independent of the type of species. A second drift velocity arises from field curvature and ∇B inhomogeneities of the magnetic field:

$$\mathbf{v}_B = \frac{m_s(v_\perp + 2v_\parallel)}{2q_s} \frac{\mathbf{B} \times \nabla B}{B^3}. \quad (2.2)$$

The subscript s denotes the dependence on the species type and it can be seen that this effect is in opposite directions for the ions and electrons. This effect is also dependent on the velocity of the particles and is therefore temperature dependent. This can give rise to charge separation and subsequently an ExB drift.

Important quantities in plasma physics are the flux surfaces but to properly introduce them we first require a description of the coordinate system used in plasma physics. The coordinate system used here is that of a cylindrical coordinate system (R, ϕ, z) . Here z gives the vertical distance on the axis of the tokamak, R gives the distance from the axis and ϕ of course gives the angle. This coordinate system is depicted in figure 2.1: The magnetic field can be decomposed into two components. One of which points parallel to the toroidal direction (\mathbf{e}_ϕ) and the other orthogonal to \mathbf{e}_ϕ (the poloidal direction). This is the following decomposition:

$$\mathbf{B} = \mathbf{e}_\phi B_\phi + \mathbf{B}_p. \quad (2.3)$$

We can relate this decomposition to two different flux functions. The first is the poloidal flux function, usually denoted by ψ , which gives the magnetic flux through the disk of radius R in the plane orthogonal to \mathbf{e}_z . This flux function is then given by:

$$\psi(R, z) = - \int \mathbf{B} \cdot d\mathbf{A}_z \quad (2.4)$$

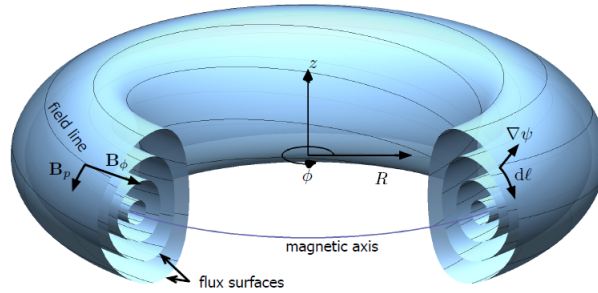


Figure 2.1: *Coordinate system of Tokamak physics*

which does not rely on ϕ because we assume the magnetic field is axis-symmetric for a tokamak. Through this definition we can related the poloidal magnetic field:

$$\mathbf{B}_p = \mathbf{e}_\phi \times \frac{\nabla\psi}{2\pi R} \quad (2.5)$$

Figure 2.1 also shows the loci of constant ψ which are called **magnetic flux surfaces** because it can be seen that $\nabla\psi \cdot \mathbf{B} = 0$ always holds for an axis-symmetric system. This implies that the magnetic field lines always line on the constant poloidal flux surfaces.

2.1 Grad-Shafranov equation

When were are looking at the force equation in equilibrium we find that the electromagnetic force needs to counter the pressure gradient force. This reads:

$$\mathbf{E} \sum_j n_j e_j + \sum_j n_j e_j \mathbf{v}_j \times \mathbf{B} = \nabla p. \quad (2.6)$$

For effects on larger length scale then the Debye length and similarly defined time scale we can impose quasi-neutrality. This will cause the first term of 2.6 to vanish and this leads to the MagnetoHydroDynamics (MHD) equilibrium:

$$\mathbf{j} \times \mathbf{B} = \nabla p. \quad (2.7)$$

This will of course lead to the following equations: $\mathbf{B} \cdot \nabla p = \mathbf{j} \cdot \nabla p = 0$. To compute 2.7 in terms of the flux function we will need to do some work. Firstly we can apply Faraday's law to 2.3 to find an expression for the current

density:

$$\mu_0 \mathbf{j} = \mu_0 j_\phi \mathbf{e}_\phi + \nabla \times (B_\phi \mathbf{e}_\phi) = \mu_0 j_\phi \mathbf{e}_\phi + \frac{1}{R} \nabla(RB_\phi) \times \mathbf{e}_\phi \quad (2.8)$$

$$\mu_0 j_\phi = \nabla \times (\mathbf{e}_\phi \times \frac{\nabla \psi}{2\pi R}) = \frac{1}{2\pi R} R^2 \nabla \cdot (\frac{\nabla \psi}{R^2}) \equiv \frac{1}{2\pi R} \Delta \psi \quad (2.9)$$

Where the operator Δ^* is slightly different then the Laplace operator in cylindrical coordinates. This relates the current density to the magnetic fields and magnetic flux. It can be easily shown that $\nabla p \parallel \nabla \psi$ and $\nabla(RB_\phi) \parallel \nabla \psi$ and thus that $p(\psi)$ and $T(\psi) = RB_\phi$ are constant on a flux surface. Now we can compute the force balance equation in terms of ψ :

$$\begin{aligned} \mu_0 \nabla p &= \mu_0 \mathbf{j} \times \mathbf{B} \\ &= -\mu_0 j_\phi \frac{\nabla \psi}{2\pi R} - \frac{\nabla T}{R} \frac{T}{R} \end{aligned} \quad (2.10)$$

$$\nabla \psi \frac{dp}{d\psi} = \frac{\nabla \psi}{2\pi R} \frac{1}{2\pi R} \Delta^* \psi + \frac{\nabla \psi}{R} \frac{dT}{d\psi} \frac{T}{R}. \quad (2.11)$$

Rewriting the last equation leads to the Grad-Shafranov equation which is an equilibrium equation for the MHD:

$$\Delta^* \psi = -4\pi^2 \left(\mu_0 R^2 \frac{dp}{d\psi} + \frac{dT}{d\psi} T \right). \quad (2.12)$$

The magnetic field and current density can be written in terms of T and ψ :

$$\mathbf{B} = \mathbf{e}_\phi \frac{T(\psi)}{R} + \mathbf{e}_\phi \times \frac{\nabla \psi}{2\pi R} \quad (2.13)$$

$$\mu_0 \mathbf{j}_0 = \frac{1}{R} \frac{dT}{d\psi} (\nabla \psi \times \mathbf{e}_\phi) + \mathbf{e}_\phi \frac{1}{2\pi R} \Delta^* \psi. \quad (2.14)$$

2.2 Flux Averaging and Poloidal Flux Quantities

As we have seen, quantities $p(\psi)$ and $T(\psi)$ are constant on the poloidal flux surfaces. Other interesting quantities might not be and it is useful to introduce the notion of flux-surface-averaging before proceeding. The beginning of this description with the calculation of the volume and poloidal area of a flux surface:

$$V = \int dV = \int R d\phi \frac{d\psi}{|\nabla \psi|} dl_p = \int d\psi \int \frac{dl_p}{B_p} \quad (2.15)$$

$$A_\phi = \int \frac{d\psi}{|\nabla \psi|} dl_p = \int d\psi \oint \frac{l_p}{2\pi R B_p} \quad (2.16)$$

where dl_p is an infinitesimal length along the poloidal flux surface perpendicular to \mathbf{e}_ϕ . Quantities that are not constant across a flux-surface can be averaged over this surface in the following way:

$$\langle Q \rangle = \frac{\partial}{\partial V} \int Q dV = \frac{\partial \psi}{\partial V} \frac{\partial}{\partial V} \int Q \frac{R dl_p}{|\nabla \psi|} d\psi d\phi = \oint Q \frac{dl_p}{B_p} / \oint \frac{dl_p}{B_p}. \quad (2.17)$$

Another important magnetic flux is that of the toroidal flux defined as:

$$\Phi = \int \mathbf{B} \cdot d\mathbf{S}_\phi = \frac{1}{2\pi} \int_V \mathbf{B} \cdot \nabla \psi dV = \frac{1}{2\pi} \int_V \frac{T}{R^2} dV \quad (2.18)$$

so that:

$$\frac{\partial \Phi}{\partial V} = \frac{1}{2\pi} \frac{\partial}{\partial V} \int_V \frac{T}{R^2} dV = \frac{1}{2\pi} \langle T/R^2 \rangle. \quad (2.19)$$

We can define a coordinate that is an effective minor radius with respect to this toroidal flux:

$$\rho = \sqrt{\frac{\Phi}{\pi B_0}} \quad (2.20)$$

To look at some flux averaged quantities it is useful to introduce some widely used notation:

$$J = \frac{T(\psi)}{R_0 B_0} \quad (2.21)$$

$$G_2 = \frac{V'}{4\pi^2} \left\langle \left(\frac{(\nabla \rho)^2}{R^2} \right) \right\rangle \quad (2.22)$$

$$V' = \frac{\partial V}{\partial \rho}. \quad (2.23)$$

The function J is the normalized poloidal current function T and it reflects the diamagnetic effect of the plasma and G_2 is a geometric quantity that depends on the flux surface configuration. Using these definitions we will calculate some important quantities:

Plasma current and current density

$$I_{pl}(\rho) = \frac{G_2}{\mu_0} \frac{\partial \psi}{\partial \rho} \quad (2.24)$$

$$j_{torr} = R_0 \left\langle \frac{\mathbf{j} \cdot \mathbf{e}_\phi}{R} \right\rangle \quad (2.25)$$

Safety Factor and Magnetic Shear

A widely used quantity related to plasma instabilities and transport are the safety factor

$$q \equiv \frac{2\pi\rho B_0}{\frac{d\psi}{d\rho}} \quad (2.26)$$

and its rate of change, referred to as the magnetic shear:

$$s = \frac{V}{q} \frac{\partial q}{\partial V} = \frac{V}{\rho V'} \left(1 - \rho \frac{\partial^2 \psi}{\partial \rho^2} \left(\frac{\partial \psi}{\partial \rho} \right)^{-1} \right) \quad (2.27)$$

Poloidal Field and Magnetic Energy

We can cast the poloidal field strength in a different form:

$$B_p = \frac{|\nabla \rho|}{2\pi R} \frac{\partial \psi}{\partial \rho}. \quad (2.28)$$

In a cylindrical poloidal flux field it would be constant on the flux surface but it is not constant on a flux surface in general. The magnetic energy density of the plasma poloidal field is given by:

$$w_i = \frac{\langle B_p^2 \rangle}{2\mu_0} = \frac{G_2}{2\mu_0 V'} \left(\frac{\partial \psi}{\partial \rho} \right)^2 \quad (2.29)$$

Integration over the volume inclosed by the flux surface ρ gives the total magnetic energy enclosed by this flux surface:

$$W_i = \frac{1}{2\mu_0} \int_0^\rho \left(\frac{\partial \psi}{\partial \rho} \right)^2 G_2 d\rho \quad (2.30)$$

2.3 Transport Equations

Very important aspect of our simulations is going to be the transport of different quantities through the plasma. These will tell us a lot about the evolution of the system. These derivations will rely on definitions and calculations in the previous section. Most important are the poloidal flux diffusion, energy transport and the particle transport because these quantities are all important to the confinement time.

2.3.1 Poloidal Flux Diffusion

The poloidal flux diffusion equation as a function of ρ is a reexpression of Ohm's law averaged over a flux surface. The derivation needs some arithmetics and intermediate steps but is insightful for the way the other diffusion

equations are calculated in the flux averaging procedure. The starting point of this calculation is Ohm's law projected parallel to the magnetic field:

$$j_{\parallel} = \sigma_{\parallel} E_{\parallel} + (j_{bs} + j_{cd}) \quad (2.31)$$

with $j_{\parallel} = \frac{\langle \mathbf{j} \cdot \mathbf{B} \rangle}{B_0}$.

Some of the identities we will use to derive the full flux diffusion equation are too detailed for this part of the thesis but they will be derived in the Appendix. Two important quantities are:

$$\langle \nabla \cdot \mathbf{F} \rangle = \frac{\partial}{\partial V} \langle \mathbf{F} \cdot \nabla V \rangle, \quad (2.32)$$

and for a scalar $H(t) = \int_V F dV$ where the enclosed volume is characterized by constant ϕ :

$$\left. \frac{\partial H}{\partial t} \right|_{\psi=\text{cst}} = \int_V \frac{\partial F}{\partial t} dV + \oint_S F \mathbf{u}_{\psi} \cdot \frac{\nabla \psi}{|\nabla \psi|} dS_{\psi} \quad (2.33)$$

Using these identities we can express the time derivative of the toroidal flux enclosed by a surface S of constant poloidal flux as:

$$\left. \frac{\partial \Phi}{\partial t} \right|_{\psi=\text{cst}} = \frac{1}{2\pi} \frac{\partial}{\partial t} \int_V \mathbf{B} \cdot \nabla \phi dV \quad (2.34)$$

$$= \frac{1}{2\pi} \int_V \frac{\partial \mathbf{B}}{\partial t} \cdot \nabla \phi dV + \frac{1}{2\pi} \oint_S (\mathbf{B} \cdot \nabla \phi) (\mathbf{u}_{\psi} \cdot \nabla \psi) \frac{dS}{|\nabla \psi|} \quad (2.35)$$

We will need this expression later on to express the time derivative of the poloidal flux. We will first want to further evaluate this expression. We can express it using the poloidal and toroidal electric field using the following identities.

$$\int_V \frac{\partial \mathbf{B}}{\partial t} \cdot \nabla \phi dV = -2\pi \oint_S \mathbf{E} \cdot \mathbf{B}_p \frac{dS}{|\nabla \phi|} \quad (2.36)$$

$$\mathbf{u}_{\psi} \cdot \nabla \psi = -2\pi R E_{\phi}. \quad (2.37)$$

Using equations in (2.36) we can reexpress (2.34):

$$\begin{aligned} \left. \frac{\partial \Phi}{\partial t} \right|_{\psi=\text{cst}} &= - \oint_S (\mathbf{E} \cdot \mathbf{B}_p + B_{\phi} E_{\phi}) \frac{dS}{|\nabla \psi|} \\ &= - \oint_S (\mathbf{E} \cdot \mathbf{B}) \frac{dS}{|\nabla \psi|} \\ &= - \frac{\partial V}{\partial \psi} \langle \mathbf{E} \cdot \mathbf{B} \rangle \end{aligned} \quad (2.38)$$

Now this expression is present in equation (2.31) as the parallel electric field. Now we wish to relate this to the time derivative of the poloidal flux:

$$\left. \frac{\partial \psi}{\partial t} \right|_{\Phi=\text{cst}} = \left. \frac{\partial \psi}{\partial t} \right|_{\rho} + \left. \frac{\partial \psi}{\partial \rho} \frac{\partial \rho}{\partial t} \right|_{\Phi} \quad (2.39)$$

$$= \left. \frac{\partial \psi}{\partial t} \right|_{\rho} - \frac{\rho B_0}{2B_0} \frac{\partial \psi}{\partial \rho} \quad (2.40)$$

but on the other hand:

$$\left. \frac{\partial \psi}{\partial t} \right|_{\Phi=\text{cst}} = \left. \frac{\partial \psi}{\partial V} \frac{\partial V}{\partial \Phi} \frac{\partial \Phi}{\partial t} \right|_{\psi=\text{cst}} \quad (2.41)$$

$$= -\frac{\partial V}{\partial \Phi} \langle \mathbf{E} \cdot \mathbf{B} \rangle \quad (2.42)$$

$$= -2\pi \frac{\langle \mathbf{E} \cdot \mathbf{B} \rangle}{T \langle 1/R^2 \rangle} \quad \text{using 2.19} \quad (2.43)$$

The last ingredient from (2.31) we still need is the following expression, derived in A.7:

$$\frac{\langle \mathbf{j} \cdot \mathbf{B} \rangle}{B_0} = \frac{2\pi R_0 J^2}{\mu_0 V'} \frac{\partial}{\partial \rho} \left(\frac{G_2}{J} \frac{\partial \psi}{\partial \rho} \right) \quad (2.44)$$

Putting everything together we arrive at the poloidal flux diffusion equation:

$$\sigma_{\parallel} \left(\frac{\partial \psi}{\partial \rho} - \frac{\rho B_0}{2B_0} \frac{\partial \psi}{\partial \rho} \right) = \frac{R_0 J^2}{\mu_0 \rho} \frac{\partial}{\partial \rho} \left(\frac{G_2}{J} \frac{\partial \psi}{\partial \rho} \right) - \frac{V'}{2\pi \rho} (j_{bs} + j_{cd}). \quad (2.45)$$

2.3.2 Particle Transport

The continuity equation for a particular plasma particle species α reads the following:

$$\frac{\partial n_{\alpha}}{\partial t} + \nabla \cdot (n_{\alpha} \mathbf{u}_{\alpha}) = s_{\alpha}, \quad (2.46)$$

where n_{α} is species α 's density, \mathbf{u}_{α} the fluid velocity of that particle species and s_{α} is the localized particle source. This expression can be reexpressed in terms of flux averaging expressions. This derivation is done in the appendix. The final result is the particle flux equation:

$$\frac{1}{V'} \left(\frac{\partial}{\partial t} - \frac{\rho B_0}{2B_0} \frac{\partial}{\partial \rho} \right) (\langle n_{\alpha} \rangle V') + \frac{1}{V'} \frac{\partial}{\partial \rho} \Gamma_{\alpha} = S_{\alpha}. \quad (2.47)$$

The particle fluxes Γ_{α} are set by various complex processes which will not be further treated in this thesis. There are certain sources of fluxes which do not lie in particle sources.

2.3.3 Energy Transport

The derivation of the energy transport is a lengthy one but is done in a similar manner as the particle transport equation (2.47). The end result of this derivation can be written in a form that is very insightful and easily displays the different contribution to energy transport in the plasma. The details of the derivation will not be discussed but this (and all other transport equations) can be found in Hinton [2]. The result:

$$\boxed{\frac{3}{2}V^{5/3}\left(\frac{\partial}{\partial t} - \frac{\rho B_0}{2B_0}\frac{\partial}{\partial \rho}\right)[V^{5/3}n_\alpha T_\alpha] + \frac{1}{V'}\frac{\partial}{\partial \rho}\left(q_\alpha + \frac{5}{2}T_\alpha \Gamma_\alpha\right) = P_\alpha} \quad (2.48)$$

Where P_α is the power supplied to the particular particle species. The heat fluxes that are dissipated convectively and diffusively are given by:

$$\Gamma_\alpha = -V'G_1n_\alpha \left[\sum_{\beta \in \text{species}} \left(D_{n_\beta}^\alpha \frac{1}{n_\beta} \frac{\partial n_\beta}{\partial \rho} + D_{T_\beta}^\alpha \frac{1}{T_\beta} \frac{\partial T_\beta}{\partial \rho} \right) + D_E \frac{E_\parallel}{B_p} \right] \quad (2.49)$$

$$q_\alpha = -V'G_1T_\alpha n_\alpha \left[\sum_{\beta \in \text{species}} \left(\chi_{n_\beta}^\alpha \frac{1}{n_\beta} \frac{\partial n_\beta}{\partial \rho} + \chi_{T_\beta}^\alpha \frac{1}{T_\beta} \frac{\partial T_\beta}{\partial \rho} \right) + \chi_E \frac{E_\parallel}{B_p} \right] \quad (2.50)$$

The coefficient G_1 is again a geometric quantity and the transport coefficients $\chi_{x_\beta}^\alpha, D_{x_\beta}^\alpha$ quantify the transport due to a gradient in a given profile. These coefficients themselves may rely on the position in the plasma and other plasma quantities such as the the profiles. Finding correct expressions for these transport coefficients and solving this equation numerically is one of the important tasks at hand.

2.4 Sources

Particle Sources

The most important sources of particles are:

- Gas injection system such as gas puffing and pellet injection
- Neutral beam injection
- particles released from the wall through collisions
- Fusion is a source of α -particles.

Whereas the following are particle sinks:

- Cryogenic pumping at the wall
- Atomic processes such as ionization of injected particles
- Fusion acts as a sink of Tritium and Deuterium

Energy

Energy sources or sinks can be:

- $P_{e,i}$ is the electron-ion equipartition power that attempts to decrease the temperature difference. This loss term is given by:

$$P_{ei} = n_e \nu_{ei} (T_e - T_i), \quad (2.51)$$

where ν_{ei} is the ion-electron relaxation time.

- P_{rad} is the radiation loss which can be divided up into bremsstrahlung, line radiation and cyclotron radiation. The bremsstrahlung can be quantified by a simple expression, taken from Wesson [1]:

$$P_{br} = 5.35 \times 10^{-37} Z_{eff} n_e^2 T_e^{1/2} \text{W m}^{-3} \quad T_e \text{ in keV.} \quad (2.52)$$

The line radiation is not so easily quantified and is different for each plasma species.

- P_{atomic} are atomic processes which cause energy losses and gains.
- P_{fusion} is a main source of energy through fusion reactions.
- P_{aux} is a combined factor for all auxiliary heating systems such as Neutral Beam Injection, Ohmic Heating and RF heating.

2.5 Heating Systems

To reach the required temperature of 10-20 keV we need to substantially heat the plasma. The initial heating in all tokamaks comes from the ohmic heating caused by the toroidal current. At low temperatures the ohmic heating is quite effective and can easily reach a temperature of a few keV. To reach higher temperatures additional heating schemes are required to reach ignition. There are two main external heating schemes available for further heating. These are the injection of energetic neutral beams and the resonant absorption of electromagnetic waves. After reaching the fusion regime the heating will be assisted by the energetic alpha particles created during the fusion process. The external heating systems used in ITER will be discussed

in more detail in the following sections.

Ohmic Heating

Ohmic heating works just as in metals or other conductive materials. When a current is applied to a system then it is heated through collisions. The ohmic heating density is given by:

$$P_{\Omega} = \eta j^2. \quad (2.53)$$

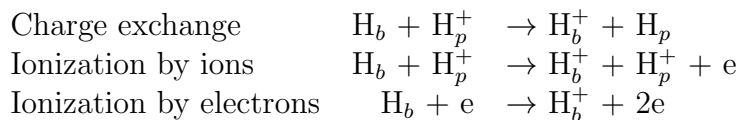
where η is the resistivity of the plasma and j the usual current density. As mentioned earlier, ohmic heating cannot account for all the heating needed to reach ignition. A problem with Ohmic heating of plasma's is that the resistivity of a plasma, in contrast to metals, decreases as the temperature rises $\eta \propto T^{-3/2}$. Normally this would be favorable behavior but in this case we cannot heat the plasma to the desired fusion temperature by solely using Ohmic heating. There are also limitations to the achievable current density imposed by plasma instabilities. The temperature achievable is dependent on the scaling law of the confinement time τ_E but a particular choice called the Alcator scaling leads to a peak temperature given by :

$$\hat{T} = 2.1 \left(\frac{Z_{eff}}{q_a q_0} \right)^{2/5} \left(\frac{a B_{\phi}}{R} \right)^{4/5} \text{ keV}. \quad (2.54)$$

For a particular configuration ($Z_{eff} = 1.5$, $q_a a_0 =$ and $R/a =$) this leads to a peak temperature of ... at magnetic field of ...

Neutral Beam Injection (NBI)

In order to inject energetic particles into the plasma they need to be neutrally charged since otherwise they would be reflected by the magnetic fields of the plasma. These neutral atoms travel into the plasma in straight lines until they become ionized through collisions with plasma particles. The injected atoms are subsequently slowed down and eventually they thermalize. This mean of heating has 3 basic effects: heating, driving current and adding fuel (when deuterium or tritium beams are used). There are 3 processes which lead to ionization of the neutral beam:



The absorption of the number of beam particles is governed by the cross-sections for these different processes. The cross-sections for these processes are shown in Figure 2.2: In a reactor the electron density will be around 10^{20}

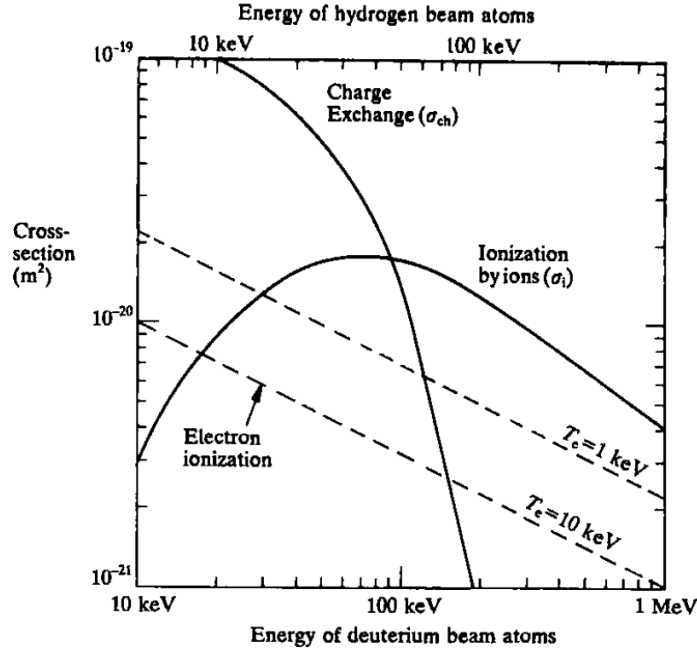


Figure 2.2: *The different ionization cross-sections for the NBI*

m^{-3} . For a deuteron beam energy of 100 keV the combined cross-sections lead to a decay length of 30 cm. The injected beam energy needs to be sufficiently high to avoid infringements on the reactor wall.

The beam ion heating is caused by a friction force between the plasma particles and the beam particles after ionization. The energy transfer is dominated by the beam ions because of their high mass and therefore high energy. The beam ions lose energy through collision with the plasma and each collision transfers energy from the beam to the plasma. Sivukhin [3] showed in 1966 that the energy decrease of a particle due to the background species s with Maxwellian distribution f_s and thermal velocity $v_{T,s}$ is dictated by:

$$\frac{dE_b}{dt} = -\frac{4\pi Z_b^2 e^2}{v_b} \sum_s (\ln \Lambda) n_s Z_s^2 e^2 \times \left[\frac{\text{Erf}(w_s)}{m_s} - \frac{2w_s(m_s + m_b)}{m_s m_b \sqrt{\pi}} e^{-w_s^2} \right] \quad (2.55)$$

where $w_s = v_b/\sqrt{2}V_{T,s}$ and the index b refers to the beam ions. To make a qualitative assessment of the energy transfer we can make the following assumption: A 100 keV ion has a velocity of $3 \cdot 10^6$ m/s whereas in a 5 keV plasma the thermal velocity of the electrons and ions are respectively $3 \cdot 10^7$ m/s and $5 \cdot 10^5$ m/s. Therefore we can simplify 2.55 by using $\text{Erf}(x) \approx 2x/\sqrt{\pi}$

for $x \ll 1$ and $\text{Erf}(x) \approx 1$ for $x \gg 1$. The result can be written as:

$$\frac{dE_b}{dt} = -\frac{2E_b}{\tau_s} \left\{ 1 + \left(\frac{E_c}{E_b} \right) \right\}, \quad (2.56)$$

where the first term corresponds to the energy transfer to the electrons and the second term to the plasma ions and E_c is the critical energy at which the energy to ions is equal to that to the electrons. The beam velocity at which this happens is called the critical velocity:

$$v_c = \sqrt{\frac{2k_b T}{m_e}} \left(\frac{3\sqrt{\pi}}{4} \sum_i n_i n_e Z_i^2 \frac{m_e}{m_i} \right)^{1/3}. \quad (2.57)$$

The fraction of the initial beam ion energy that is transferred to the plasma ions can be calculated by using that:

$$P_i = -\frac{E_c^{3/2}}{E_c^{3/2} + E_b^{3/2}} \frac{dE_b}{dt} \quad (2.58)$$

$$\int P_i dt = E_c^{3/2} \int_0^{E_{b0}} \frac{dE_b}{E_c^{3/2} + E_b^{3/2}} \quad (2.59)$$

$$\frac{1}{E_{b0}} \int P_i dt = \frac{E_c}{E_{b0}} \int_0^{E_{b0}/E_c} \frac{E_c d(E_b/E_c)}{1 + (E_b/E_c)^{3/2}} \quad (2.60)$$

where the last integral is known.

Radio Frequency Heating

The other important heating scheme is based on the resonant absorption of radio frequency electromagnetic waves which is a collisionless process and can produce strong heating. A magnetized, multi species plasma has multiple resonant frequencies that give rise to absorption. This variety in wave modes allows for many different radio frequency heating schemes, all of which consist of the same general layout. The cold plasma dispersion relation gives a good approximation to the propagation of electromagnetic waves in a plasma. This model can therefore be used to investigate the three principal radio frequency heating schemes. The cold plasma dispersion relation is give by:

$$\epsilon_{\perp} n_{\perp}^4 - [(\epsilon_{\perp} - n_{\parallel}^2)(\epsilon_{\perp} + \epsilon_{\parallel}) + \epsilon_{xy}^2] n_{\perp}^2 + \epsilon_{\parallel} [(\epsilon_{\perp} - n_{\parallel}^2)^2 + \epsilon_{xy}^2] = 0 \quad (2.61)$$

where x and y represent the radial and poloidal direction and the z -direction is the equilibrium magnetic field direction. All \perp and \parallel direction are with

respect to the equilibrium magnetic field direction and n is the refractive index. The dielectric tensor has the following components:

$$\epsilon_{xx} = \epsilon_{yy} = \epsilon_{\perp} = 1 - \sum_j \frac{\omega_{pj}^2}{\omega^2 - \omega_{cj}^2}, \quad (2.62)$$

$$\epsilon_{zz} = \epsilon_{\parallel} = \sum_j \frac{\omega_{pj}^2}{\omega^2}, \quad (2.63)$$

$$\epsilon_{xy} = \sum_j \frac{i\omega_{pj}^2\omega_{cj}^2}{\omega(\omega_{cj}^2 - \omega^2)} \text{ and} \quad (2.64)$$

$$\epsilon_{yx} = -\epsilon_{xy} \quad (2.65)$$

Where the summation is over electrons and all ion species, ω_{cj} is the gyrofrequency of species j and ω_{pj} is the plasma frequency of species j . Since (2.61) is quadratic in n_{\perp}^2 , one of the solution will tend to infinity when the coefficient of n_{\perp}^4 becomes zero. So $\epsilon_{\perp} = 0$ is the condition for a perpendicular cold plasma resonance. There are three main heating scenarios using radio frequency absorption. These will be briefly discussed in the order of their resonance frequencies:

- **Ion Cyclotron Heating (ICRH)**

The lowest frequency resonance scheme is the *ion cyclotron resonance heating* for frequencies $\omega \sim \omega_{ci}$. This ion resonance has the limitation that it needs two (or more) ion species present in the plasma. The resonant frequency ω_{ii} with two ion species is:

$$\omega_{ii}^2 = \frac{\omega_{c1}\omega_{c2}(1 + n_2m_2/n_1m_1)}{m_2Z_1/m_1Z_2 + n_2Z_2/n_1Z_1} \quad (2.66)$$

and it gives a resonant frequency $\sim 40\text{MHz}$.

- **Lower Hybrid Heating and Current Drive (LHCD)**

The *lower hybrid resonance* frequency lies between ω_{ci} and ω_{ce} and relies on resonant coupling to a wave in the plasma. For central regions of a tokamak and neglecting terms $\sim m_e/m_i$ the lower hybrid resonance is given by:

$$\omega_{LH}^2 \approx \frac{\omega_{pi}^2}{1 + \omega_{pi}^2/\omega_{ce}^2}. \quad (2.67)$$

This resonance lies within the range $\sim 1\text{-}8\text{ GHz}$ and for a fusion plasma is typically around 5 GHz .

- **Electron Cyclotron Heating and Current Drive (ECRH/ECCD)**

The highest frequency scheme is the *electron cyclotron resonance* for frequencies $\omega \sim \omega_{ce}$ which couples to the electron cyclotron frequency. The resonance frequency is given by the upper hybrid frequency:

$$\omega_{UH}^2 = \omega_{pe}^2 + \omega_{ce}^2 \quad (2.68)$$

This is just above the electron cyclotron frequency and is in the range 100-200 GHz. The disadvantage of these last two heating schemes is that they heat the electrons instead of the more favorable ions.

2.6 Fusion Reaction

As most people know the most easily accessible fusion reaction is that between deuterium and tritium. This reaction produces an α -particle with 3.51 MeV and a neutron with 14.1 MeV that stems from the mass deficit of 17.6 MeV. This reaction is induced through collisions between the particles in the plasma. The Coulomb barrier prevents the fusion reaction to take place at low impact energies. The D-T fusion reaction is favorable because of quantum mechanical tunneling and occurs at energies just below that needed to overcome the Coulomb barrier. The cross section for this and other hydrogen based fusion reactions is given in figure ??:

The reaction rate of the D-T fusion reaction can be calculated using the distribution functions of the two species:

$$\mathcal{R} = \int \int \sigma(v') v' f_1(\mathbf{v}_1) f_2(\mathbf{v}_2) d^3 v_1 d^3 v_2 \quad (2.69)$$

where v' is the relative velocity between \mathbf{v}_1 and \mathbf{v}_2 . We know that for the ions in the plasma that are in thermal equilibrium we can write their distribution function as:

$$f_i(v_i) = n_i \left(\frac{m_i}{2\pi T} \right)^{3/2} \exp^{-\frac{m_i v_i^2}{2T}} \quad (2.70)$$

so that the total reaction rate per unit volume after some computation becomes:

$$\mathcal{R} = \left(\frac{8}{\pi} \right)^{1/2} n_1 n_2 \left(\frac{\mu}{T} \right)^{3/2} \frac{1}{m_1^2} \int \sigma(\epsilon) \epsilon \exp \left(-\frac{\mu \epsilon}{m_1 T} \right) d\epsilon. \quad (2.71)$$

where μ is the reduced mass and ϵ the kinetic energy. Now we need to have the cross-section of the fusion event to be able to calculate the total reaction

Coefficient	T(d,n) ⁴ He
$B_G(\text{keV})$	34.3827
$m_r c^2(\text{keV})$	1 124 656
C_1	1.17302×10^{-9}
C_2	1.51361×10^{-2}
C_3	7.51886×10^{-2}
C_4	4.60643×10^{-3}
C_5	1.35000×10^{-2}
C_6	-1.06750×10^{-4}
C_7	1.36600×10^{-5}

Table 2.1: The coefficients for the reactivity as given in equation (2.72)

rate. This cross-section was measured for different fusion processes by Bosch *et al.*[4]. The model of the reactivity used in this paper is given by:

$$\langle \sigma v \rangle = C_1 \cdot \theta \sqrt{\xi / (m_r c^2 T_i^3)} \exp(-3\xi), \quad (2.72)$$

$$\theta = \frac{T_i}{1 - \frac{T_i(C_2 + T_i(C_4 + T_i C_6))}{1 + T_i(C_3 + T_i(C_5 + T_i C_7))}}, \quad (2.73)$$

$$\xi = \left(\frac{B_G^2}{4\theta} \right)^{1/3}, \quad (2.74)$$

where $\langle \sigma v \rangle$ is in cm^3/s and T is the ion temperature in keV. The original results for the fitting of these coefficients are given in table 2.1:

The maximal error in the fitting of these parameter was .25% within $0.2 \leq T(\text{keV}) \leq 100$. It can be seen by calculating the reaction rate for an interesting temperature that the fusion reactions come predominantly from the tail of the Maxwellian distribution. Thus if we are given a density and temperature profile we can calculate, using equation (2.71), what the birth profile of the fusion particles is.

The energy generated by the fusion reaction is transported for roughly 80% by the neutrons and 20% by the α -particles. The neutrons have a very low cross section with the plasma constituents and it is save to say they leave the plasma without any collisions with the plasma and thus with 80% of the energy released by the fusion reaction. The α -particles however do have a substantial cross-section and will therefore collide with the plasma constituents. This will result in energy transfer from α -particles to the plasma

particle which in turn results in heating of the plasma. Similar equations as for the NBI hold for the energy deposition of the α -particles but here we are also interested in the distribution function of the α -particles in order to see where they deposit their energy. The starting point of this computation is the Fokker-Planck equation:

$$\frac{df_\alpha}{dt} = \sum_s C(f_\alpha, f_s) + C(f_\alpha, f_\alpha) + S + L, \quad (2.75)$$

where the S and L represent a source and lost term, the sum over s implies all plasma constituents and $C(f_\alpha, f_\alpha)$ take into account the self collisions. The derivation of this equation can be found in [3]. The collision term between plasma particles and the α -particles can be expressed using the Landau form of the collision integral:

$$C(f_\alpha, f_s) = \frac{Z_\alpha^2 Z_s^2 e^4 \ln(\Lambda)}{8\pi\epsilon_0^2 m_\alpha} \frac{\partial}{\partial \mathbf{v}} \cdot \int \frac{u^2 \mathbf{I} - \mathbf{u}\mathbf{u}}{u^3} \left[\frac{f_s(\mathbf{w})}{m_\alpha} \frac{\partial f_\alpha(\mathbf{v})}{\partial \mathbf{v}} - \frac{f_\alpha(\mathbf{v})}{m_s} \frac{\partial f_s(\mathbf{w})}{\partial \mathbf{w}} \right] d^3 w \quad (2.76)$$

Solving equation (2.75) is a very complex problem[5] that is even heavy to evaluate numerically. However a couple of simple assumptions can be made that greatly simplify the problem. This equation becomes non-linear by the self-collisional term $C(f_\alpha, f_\alpha)$. This term can be neglected because of the relatively low density compared to the plasma ions. Another assumption that is used is that the plasma particles are isotropically Maxwellian. In this case the collision integrals can be performed analytically and the result depend on expressions of the error function. Further simplification and computation is available when the assumption previously made in the NBI section is also used here: $v_{Ti} \ll v_\alpha \ll v_{Te}$. In the case of the D-T reaction this assumption is less valid since $v_{\alpha 0} \sim v_{Te}$. In addition to this there is the fact that during the thermalization of the α -particles the velocity of these particles is reduced and finally reaches a velocity that is of the order of magnitude of v_{Ti} . When these problems are overlooked and computation is proceeded it is possible to derive the following collision operator:

$$C_1(f_\alpha) = \frac{1}{\tau_s v^3} \left[v \frac{\partial}{\partial v} [(v^3 + v_c^3) f_\alpha] + Z_2 \frac{v_c^3}{2} \frac{\partial}{\partial \mu} [(1 - \mu^2) \frac{\partial f_\alpha}{\partial \mu}] \right], \quad (2.77)$$

With $Z_2 = \frac{\sum_i n_i Z_i / m_\alpha}{\sum_i n_i Z_i / m_i}$, τ_s is the characteristic stopping time, v_c the critical velocity defined in the NBI section (2.57) and $\mu = \frac{v_\parallel}{v}$. The resulting simplified

Fokker-Planck equation can be solved for the following 2 source types:

$$S(v, \mu) = \frac{S_0}{v^2} \delta(v - v_0) \delta(\mu - \mu_0) \quad (2.78)$$

$$S(v) = \frac{S_0}{v^2} \delta(v - v_0) \quad (2.79)$$

In the case of α -particles produced by fusion reactions the source terms (2.79) is a suitable source term. Even though the produced α -particles will not be represented by a delta function, the spread in their velocity is so small it can be effectively modeled by this source term. The resulting steady state distribution function is given by:

$$f(v) = \tau_s S_0 \frac{1}{v^3 + v_c^3} \Theta(v_0 - v) \quad (2.80)$$

where $\Theta(x)$ is the heaviside step function. This distribution function is obviously very different from the Maxwellian distribution.

Chapter 3

Theory of Optimization

Optimization can be found in all sorts of scientific disciplines: from economics to engineering and manufacturing to Mother Nature. Optimization is an important tool in the analysis of physical systems. In able to make use of this tool we need to identify an *objective* which need optimization. This is a quantitative measure of the behavior of the system. This objective is subject to change under certain characteristics of the system. The goal of optimization is to find the variables that optimize the objective. Generally speaking model of the system at hand has to be expressed. The model needs to describe the characteristics and evolution of the system being analyzed. In general this cannot be done with absolute detail. In this case the model should be chosen such that it mimics the important characteristics but it should not become to overly complex. This is however a delicate business: make the model to simplistic and your answers will not give useful insights but make it to complex and the problem will be unsolvable in the time available. After the formulation of the objective and modeling of the problem the optimization can begin. The starting point of the actual calculation is choosing the appropriate algorithm. There is no universal for an optimization problem but a choose has to be made from a collection of different type of algorithms. These choose of the algorithm is up to the user and is an important one. This decision should be based on the type of problem and solution that is being investigated. General optimization problems can be classified according to the properties of the objective and constraints (convexity, linearity), the number of variables (large or small) or the smoothness of the problem (differentiable,non-differentiable). Different subsets of optimization problems can be identified:

- Discrete vs Continuous:
Some problems rely on variables that can only take values from the set

of integers whereas other problems can be based on variables that can belong to a set that is uncountable big. Operations such as permutations of and ordered set are common in discrete optimization problems but are not needed in continuous optimization problems. Another difference is the fact that depending on the smoothness of the continuous optimization problem there is a level to which information in a certain point can be related to nearby point or even give information about the global problem. This is different from discrete optimization problems where the behavior of the objective and the variables can significantly change even if two points are 'close' with regard to some measure.

- **Constrained vs Unconstrained:**
Some of the variables or the objective can be constrained by considerations of the model. This leads to a classification between optimization problems. Those that have constraints on the variables at hand and those that do not. Some times it is even safe to discard constraint for problems which would naturally possess constraints. This can be done when the algorithm and the solution to the problem are unaffected by lifting the constraint. The collection of these problems and those that are unconstrained are dealt by unconstrained optimization algorithms. These are in general more easy to solve than constrained optimization problems.
- **Linear vs Non-linear:**
This classification is characterized by the nature of the objective function and its constraints. When all of these are linear functions of the variables the problem falls in the linear-programming category. When either the objective function or the constraints are of higher order in any of the variables the problem falls in the non-linear programming category.
- **Global vs Local:**
Many of the non-linear programming algorithms are restricted to finding only a local solution of the optimization problem. This means that the solution that is found yields a smaller value of the objective function than all nearby feasible points. Some problems need to find the global solution, which is the lowest value of the objective function of all feasible points, but in most problems the global solution is hard to recognize and even more difficult to locate. For problems that are convex or linear it can be shown that the solutions obtained by looking for local solutions are the global solutions. In general however, solutions to non-linear constrained or unconstrained problems may possess local solutions

that are not global solutions.

This chapter will investigate the mathematics of optimization algorithms that are applicable to the problem of this thesis. 3 different algorithms are being looked at and for one of these an important generalization is being examined[6].

3.1 Direct Search Method

One of the most popular algorithms for unconstrained minimization of multidimensional problems is the Nelder-Mead algorithm[7], first published in 1965. It is a routine that is widely used in popular programs such as Matlab. The Nelder-Mead algorithm at each step define a nondegenerate simplex which is a geometrical figure in n dimensions with nonzero volume. At each iteration the algorithm computes one or more test points and their associated function values. The algorithm works with four different parameters that are specified before hand. These are the coefficients of *reflection*(ρ), *expansion*(χ), *contraction*(γ) and *shrinkage*(σ). These parameters should satisfy the following relations:

$$\rho > 0, \quad \chi > 1, \quad \chi > \rho, \quad 0 < \gamma < 1 \quad \text{and} \quad 0 < \sigma < 1. \quad (3.1)$$

A particular and widely used choice for these parameters are:

$$\rho = 1, \quad \chi = 2, \quad \gamma = \frac{1}{2} \quad \text{and} \quad \sigma = \frac{1}{2}. \quad (3.2)$$

To describe the algorithm it is best to walk through the steps of the algorithm:

1. **Order**

Order the n+1 vertices according to their function value: $f(x_1) \leq f(x_2) \leq \dots \leq f(x_{n+1})$

2. **Reflect**

Compute the reflection point x_r from

$$x_r = \hat{x} + \rho(\hat{x} - x_{n+1}) \quad (3.3)$$

where $\hat{x} = \sum_{i=1}^n x_i/n$ is the centroid of the n best points.

Evaluate $f_r = f(x_r)$.

If $f_1 \leq f_r < f_n$ accept the reflected point and terminate the iteration.

3. Expand

If $f_r < f_1$, calculate the expansion point x_e from:

$$x_e = \hat{x} + \chi(x_r - \hat{x}) = (1 + \rho\chi)\hat{x} - \rho\chi x_{n+1}. \quad (3.4)$$

and evaluate f_e . If $f_e < f_r$, accept x_e and terminate the iteration. Otherwise accept x_r and terminate the iteration.

4. Contract

If $f_r \geq f_n$, perform a contraction between \hat{x} and the better of x_{n+1} and x_r .

Outside

If $f_n \leq f_r < f_{n+1}$ perform an outside contraction:

$$x_c = \hat{x} + \gamma(x_r - \hat{x}) = (1 + \rho\gamma)\hat{x} - \rho\gamma x_{n+1}. \quad (3.5)$$

and evaluate f_c . If $f_c \leq f_r$ accept x_c and terminate iteration. Otherwise perform step 5.

Inside

If $f_r \geq f_{n+1}$ perform an inside contraction:

$$x_{cc} = \hat{x} - \gamma(\hat{x} - x_{n+1}) = (1 - \gamma)\hat{x} + \gamma x_{n+1}. \quad (3.6)$$

and evaluate f_{cc} . If $f_c \leq f_{n+1}$ accept x_{cc} and terminate iteration. Otherwise perform step 5.

5. Shrink

Perform a shrink of all but the best point:

$$v_i = x_1 + \sigma(x_i - x_1) \quad (3.7)$$

where $i = 2, \dots, n + 1$. The unordered vertices of the next simplex are x_1, v_2, \dots, v_{n+1} .

3.2 Line Search Method

One way of iteratively looking for the solution of a optimization method is by searching for a direction that lowers the value of the objective function at each iteration. The iteration is given by:

$$x_{k+1} = x_k + \alpha_k p_k \quad (3.8)$$

where α_k is a positive scalar that represents the length of the step taken. The length of the step is decided upon at each iteration on the basis of necessary

descend conditions. The search direction p_k is generally a vector with the same dimensionality as the variable of the problem at hand x_k . It is often written in the form:

$$p_k = -B_k^{-1}\nabla f_k \quad (3.9)$$

Here B_k is a symmetric nonsingular matrix. Different descent methods adopt different matrices. In the steepest descent method this is the unit matrix, while in Newton's method B_k is the exact Hessian $\nabla^2 f_k$. In quasi Newton methods B_k is a local approximation of the Hessian updated at each iteration. Using this definition we can show that p_k is indeed a descent direction:

$$p_k^T \nabla f_k = -(\nabla f_k)^T B_k^{-1} \nabla f_k < 0. \quad (3.10)$$

As mentioned before, different implementations of the algorithm can be chosen. This affects the choice of the search direction p_k and the step length α_k .

3.2.1 Step Length

When computing the step length α_k there is an inherent trade off choice to be made. It would be advantageous to choose α_k such that is the global minimizer of the univariate function $\phi(\cdot)$:

$$\phi(\alpha) = f(x_k + \alpha p_k), \quad \alpha > 0. \quad (3.11)$$

but in general this computation is far too expensive. A more practical strategy would be to perform an inexact line search and identify a value of α which adequately reduces the objective function f in order to reach convergence at an acceptable rate. Sophisticated line search methods can be quite complex and their entire description is beyond the scope of this thesis but an interested reader is referred to a book by Nocedal and Wright[6] on which the main lines of this chapter is based.

One of these more practical line search methods would require some necessary conditions on the step length α and terminate the search for an adequate step length once these conditions are met. The following two sets of conditions are most used:

Wolfe Conditions:

The first important condition realizes a sufficient descent in the objective function f and this is measured by the following inequality condition:

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k, \quad (3.12)$$

for some constant $c_1 \in (0,1)$ that is chosen by the user. In practice the constant c_1 is chosen to be quite small: $c_1 \sim 10^{-4}$. The sufficient decrease condition is illustrated in figure 3.1:

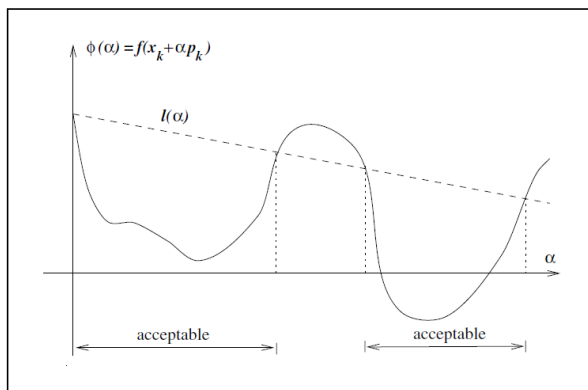


Figure 3.1: *The sufficient descent condition (Eq 3.12) illustrated*

The sufficient descent condition alone is not enough to ensure the algorithm converges at an acceptable rate. The step length should also be chosen such that the new point is sufficiently far away from our initial point. To rule out such insufficiently short step length a second requirement called the curvature condition is needed:

$$\phi'(\alpha_k) = \nabla f(x_k + \alpha_k p_k)^T p_k \geq c_2 \nabla f_k^T p_k, \quad (3.13)$$

for some constant $c_2 \in (c_1, 1)$. This condition ensures that the derivative $\phi'(\alpha_k)$ is greater than $c_2 \phi'(0)$. This condition can also be intuitively understood because when the slope $\phi'(\alpha_k)$ is strongly negative we can significantly reduce f by moving further whereas if $\phi'(\alpha_k)$ is only slightly negative we do not gain much by moving allow and we can terminate the search. This condition is illustrated in figure 3.2. Typical values of c_2 range from 0.9 until 0.1 depending on the type of line search method employed.

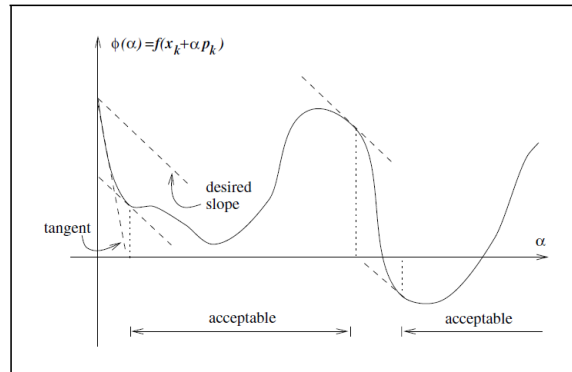


Figure 3.2: *The curvature condition (Eq 3.13) illustrated*

These two conditions make up the well known Wolfe conditions. A step length may satisfy the Wolfe conditions without being particularly close to a minimizer of ϕ . We can, however, modify the curvature condition to force α_k to lie in at least a broad neighborhood of a local minimizer or stationary point of ϕ . The strong Wolfe conditions require α_k to satisfy:

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k \quad (3.14)$$

$$|\nabla f(x_k + \alpha_k p_k)^T p_k| \leq c_2 |\nabla f_k^T p_k| \quad (3.15)$$

The only difference with the Wolfe conditions is that we no longer allow the derivative $\phi'(\alpha_k)$ to be too positive. Hence, we exclude points that are far from stationary points of $\phi(\alpha)$.

Goldstein Conditions:

A similar set of conditions are the Goldstein conditions. This set of conditions also ensure the step length α_k achieves sufficient decrease but is not too short. They are given by:

$$f(x_k) + (1 - c) \alpha_k \nabla f_k^T p_k \leq f(x_k + \alpha_k p_k) \leq f(x_k) + c \alpha_k \nabla f_k^T p_k, \quad (3.16)$$

where c can again be chosen by the user and lies in the domain $c \in (0, 1/2)$. The second inequality 3.16 again ensure sufficient decrease whereas the first inequality now ensure the length of the step from below. A disadvantage of the Goldstein conditions with respect to the Wolfe conditions is that there are possible scenario's where the first inequality rules out all possible minimizers of $\phi(\alpha)$. This set of conditions is depicted in figure 3.3:

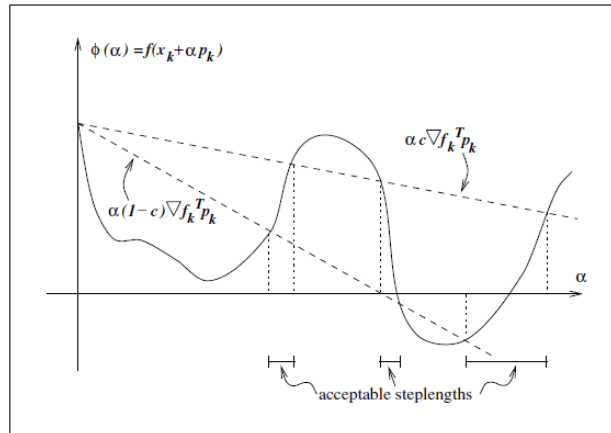


Figure 3.3: The Goldstein conditions (Eq 3.16) illustrated

3.2.2 Convergence of Line Search Methods

In order to have convergence of the algorithm to a minimizer of $\phi(\alpha)$ we need to not only choose the step length α_k appropriately but also the line search direction p_k . The angle between p_k and the steepest descent direction $-\nabla f_k$ needs to meet certain requirements in order to reach convergence of the algorithm. The following theorem by Zoutendijk quantifies the effect of appropriately chosen step lengths and can be used to show requirements on p_k so that the algorithm converges. The theorem is as follows but the proof is omitted:

Zoutendijk Theorem:

Consider any iteration of the form (3.8), where p_k is a descent direction and k satisfies the Wolfe conditions (3.13 & 3.12). Suppose that f is bounded below in \mathbb{R}^n and that f is continuously differentiable in an open set \mathcal{N} containing the level set $\mathcal{L} \equiv x : f(x) \leq f(x_0)$, where x_0 is the starting point of the iteration. Assume also that the gradient f is Lipschitz continuous on \mathcal{N} . Then

$$\sum_{k \geq 0} \cos^2 \theta_k \|\nabla f_k\|^2 < \infty. \quad (3.17)$$

The Zoutendijk condition (3.17) implies that

$$\lim_{k \rightarrow \infty} \cos^2 \theta_k \|\nabla f_k\|^2 = 0. \quad (3.18)$$

and this limit can be used to derive what conditions on p_k are needed to reach convergence. If the algorithm ensures that $\theta_k \in (-\frac{\pi}{2}, \frac{\pi}{2})$ then there is

a positive constant δ such that:

$$\cos \theta_k \geq \delta > 0 \quad \text{for all } k. \quad (3.19)$$

Then it naturally follows by combining this with (3.18) that:

$$\lim_{k \rightarrow \infty} \|\nabla f_k\|^2 = 0. \quad (3.20)$$

This shows that the algorithm that picks the direction of p_k within the appropriate direction with respect to $-\nabla f_k$ ensures that $\|\nabla f_k\|$ converges to zero, provided that it obeys some form of the Wolfe or Goldstein conditions. Equation (3.18) only ensures that the algorithm is attracted to a stationary point of the objective function. More sophisticated generalization of the line search method can ensure that the convergence reaches a local minimum. The rate of convergence is of course an important aspect of the computational speed of the algorithm. As an example of the rate of convergence we take here an ideal objective function that is a strongly convex quadratic function (3.21) and employ the steepest descent method.

$$f(x) = \frac{1}{2}x^T Qx - b^T x. \quad (3.21)$$

The error norm, which quantifies the distance between a certain iteration point and the optimum value x_0 , can be approximated for each iteration by:

$$\|x_{k+1} - x_0\|_Q^2 \leq \left(\frac{\lambda_n - \lambda_0}{\lambda_n + \lambda_0} \right)^2 \|x_k - x_0\|_Q^2. \quad (3.22)$$

Here λ_n and λ_0 are resp. the largest/smallest eigenvalue of Q and the approximation provides a worst case bound. The approximation (3.22) does however give an accurate indication of the steepest descent behavior when $n \gg 2$. The rate-of-convergence behavior for non-linear objective functions is essentially the same.

3.3 Sequential Quadratic Programming

One of the most widely used algorithm for small and large scale non-linear constrained optimization problems is that of the sequential quadratic programming (SQP). Also the optimization work presented in this thesis is based largely on this optimization scheme. The SQP scheme is an algorithm that generates iterative steps by solving quadratic programming (QP) subproblems. The SQP has a significant advantage when dealing with significant nonlinearities in the constraints of the problem. The SQP problem in its most general form is given by:

$$\min f(x) \quad (3.23)$$

$$\text{subject to } c_i(x) = 0, \quad i \in \mathcal{E} \quad (3.24)$$

$$c_j(x) \geq 0, \quad j \in \mathcal{I} \quad (3.25)$$

where \mathcal{I} and \mathcal{E} represent the collection of inequality resp. equality constraints. The functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are smooth. One can see that the constraints can be linearized by:

$$\min_p f_k + \nabla f_k^T p + \frac{1}{2} p^T \nabla_{xx}^2 \mathcal{L}_k p \quad (3.26)$$

$$\text{subject to } \nabla c_i(x_k)^T p + c_i(x_k) = 0, \quad i \in \mathcal{E}, \quad (3.27)$$

$$\nabla c_j(x_k)^T p + c_j(x_k) \geq 0, \quad j \in \mathcal{I}. \quad (3.28)$$

Let us now, for the sake of clarity, abandon the inequality constraints and replace $c_i(x)$ by $c(x)$. The lagrangian of this problem is then given by $\mathcal{L}(x, \lambda) = f(x) - \lambda^T c(x)$. The matrix A denotes the Jacobian matrix of the constraints:

$$A(x)^T = [\nabla c_1(x), \nabla c_2(x), \dots, \nabla c_m(x)] \quad (3.29)$$

Under the following assumptions:

- The constraint Jacobian $A(x)$ has full row rank
- The matrix $\nabla_{xx}^2 \mathcal{L}(x, \lambda)$ is positive definite on the tangent space of the constraints

problem 3.26 has the unique solution (p_k, λ_k) that satisfies:

$$\nabla_{xx}^2 \mathcal{L}_k p_k + \nabla f_k - A_k^T \lambda_k = 0 \quad (3.30)$$

$$A_k p_k + c_k = 0. \quad (3.31)$$

This problem can be identified as the Quadratic Programming or Newton's method problem:

$$\begin{bmatrix} \nabla_{xx}^2 \mathcal{L}_k & -A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} p_k \\ \lambda_k \end{bmatrix} = \begin{bmatrix} -\nabla f_k \\ -c_k \end{bmatrix}. \quad (3.32)$$

By restatement of the variables this can be made into a symmetric problem. The algorithm needed to solve this problem is beyond the scope of this thesis. The solution to this QP problem can be derived through many different methods such as symmetric indefinite factorization, the Schur-complement Method or the Null-space method. Explaining these methods in detail is also not very insightful but the second method shall be addressed rather briefly. The Schur-complement method is an useful way of solving the system of (3.32) when one of the following conditions are met:

- $\nabla_{xx}^2 \mathcal{L}_k$ is well conditioned and easy to convert
- $\nabla_{xx}^2 \mathcal{L}_k^{-1}$ is known explicitly through a quasi-Newton updating formula
- the number of equality constraints is small

Through a procedure called Gaussian elimination it is then possible to derive the block upper triangular system:

$$\begin{bmatrix} \nabla_{xx}^2 \mathcal{L}_k & A_k^T \\ 0 & -A_k(\nabla_{xx}^2 \mathcal{L}_k)^{-1}A_k^T \end{bmatrix}. \quad (3.33)$$

Through this procedure it is then possible to derive an explicit formula for the inverse of the symmetric form of equation (3.32):

$$\begin{bmatrix} \nabla_{xx}^2 \mathcal{L}_k & A_k^T \\ A_k & 0 \end{bmatrix}^{-1} = \begin{bmatrix} C & E \\ E^T & F \end{bmatrix}. \quad (3.34)$$

where

$$C = (\nabla_{xx}^2 \mathcal{L}_k)^{-1} - (\nabla_{xx}^2 \mathcal{L}_k)^{-1}A^T(A(\nabla_{xx}^2 \mathcal{L}_k)^{-1}A^T)^{-1}A(\nabla_{xx}^2 \mathcal{L}_k)^{-1}, \quad (3.35)$$

$$E = (\nabla_{xx}^2 \mathcal{L}_k)^{-1}A^T(A(\nabla_{xx}^2 \mathcal{L}_k)^{-1}A^T)^{-1} \quad (3.36)$$

$$F = -(\nabla_{xx}^2 \mathcal{L}_k)^{-1}A^T(A(\nabla_{xx}^2 \mathcal{L}_k)^{-1}A^T)^{-1}. \quad (3.37)$$

Chapter 4

Alpha Particle Heating

A part of the work on this thesis was to develop a model for the alpha particle heating to the electrons. This process, as mentioned before, is complex and complicated. To describe the entire process would take considerable computation time but a simplification might suffice for RAPTOR. This assumption was tried and tested within the work of this thesis. This chapter will describe the models considered for implementation and the results from these models with respect to the full model implemented in the CRONOS simulation code. First of all it is very important to identify which assumptions made in the theory section are violated and identify how much effect these violations have on the calculated deposition profiles. The main problem is the fact that after the alpha particles are created in a fusion event, the energy deposition to the plasma constituents is not instantaneous. During the deposition process, the fast alpha particles diffuse through the plasma, away from their 'birth' magnetic surface. This causes the problem that the deposition process is not completely described by the theory as described in the earlier section.

To completely describe the deposition and diffusion of the fast alpha particles considerable computation would be needed to fully solve the Focker-Planck equation in Eq (2.75) using Monte-Carlo simulations. This is implemented in the CRONOS simulation code used for ITER-like discharge simulations in a separate module called SPOT but this method is too computationally demanding to be feasible for RAPTOR. Therefore it was needed to find a simple model that is less demanding in terms of computation time. To this end, two different models were developed and tested against the SPOT model. In the following sections the specifications of these models are discussed and the results of the benchmark are shown.

4.1 Direct Deposition Model

The first model is also the least complex and operates under the assumption that the characteristic stopping time is long compared to plasma collision times and short compared to the simulation time step. Also it is assumed that the stopping time is several orders of magnitude smaller than the fast ions and plasma constituents diffusion timescale. This way we may assume that the diffusion of the fast ions plays a negligible role on the energy deposition profile. These requirements lead to an estimate for the stopping time between . . . s and . . . s for ITER. When this requirement is met the deposition profile can be approximated by using the birth profile of the alpha particles. The procedure to calculate the deposition profile is then straightforward: Given the ion and electron temperature and density profiles, the birth profile (Eq (2.72)) and the energy fraction to electrons/ions (Eq (2.60)) can be calculated. Combining these results with the difference of initial and final alpha particle energy gives the energy deposition profile to the electrons and ions as the energy created by the fusion event at a certain position in the plasma. Thus this model approximates the problem by completely neglecting any transport of fast ions during the energy transfer.

4.2 Simple Diffusion Model

The second model that was considered does, to some degree and in a specific way, take into account the diffusion of the fast alpha particles during the energy transfer. As mentioned before, solving the diffusion equation is a difficult and demanding task. As an example: keeping track of the position, collisions and energy of the fast ions during their thermalization is a tedious and computationally demanding job. For the model used here, however, a different approach was used. To introduce the diffusion of the fast ions we first turned to the well known result for the source free 1D diffusion equation. This result is given by equation (4.1):

$$n(x, t) = n_0 \frac{e^{-\frac{x^2}{4Dt}}}{2\sqrt{\pi Dt}} \quad (4.1)$$

In the simple diffusion model this result is used at each space point of the space discretized grid to calculate a diffusion term. This is done by using a diffusion coefficient taken from literature [8] ($D = 0.1m^2/s$) and calculating the alpha particle density and stopping time. These are then used to calculate the density contribution from diffusion at each grid point. This model then gives the result from a simplistic diffusion model for the birth

density profile of fast alpha particles and should therefore give a better result for the energy deposition profile. It is however a more complex calculation and when the direct and simple diffusion model do not differ all that much the direct approach will be preferable for RAPTOR. Especially since the RAPTOR code also uses analytic first derivatives which are in the case of the simple diffusion model much more complex. Figure ?? shows the simple diffusion model contributions as a function of ρ :

This model can be used to describe a diffusion with respect to nearest neighbors or to any degree of neighbors. In the work of this thesis the effect of introducing higher orders of neighbors was also analyzed. The effects on the fast ions particle flux are shown in figure 4.1 and it can be seen that the higher orders correspond to bigger particle flux:

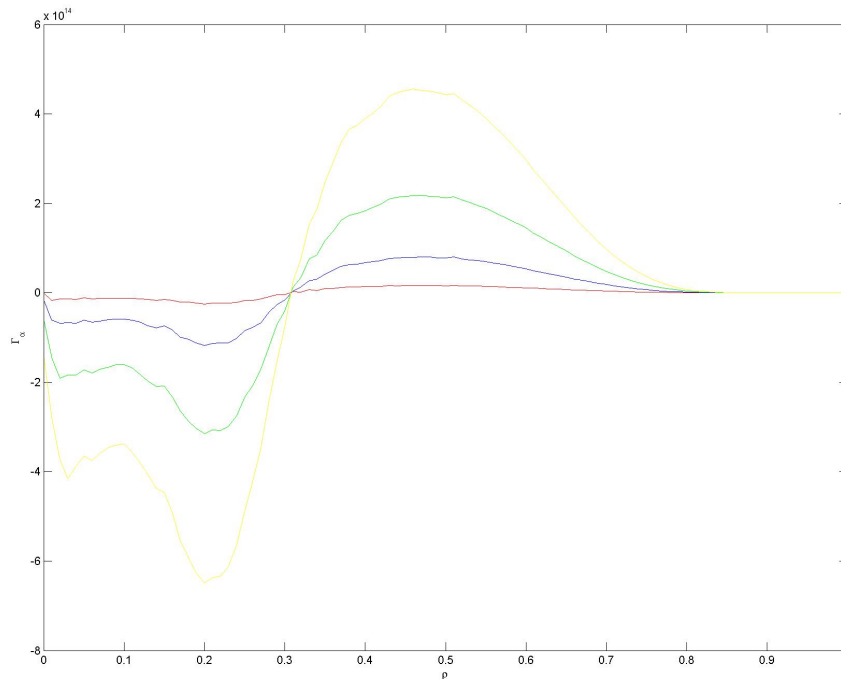


Figure 4.1: *The effects of introducing higher orders of nearest neighbors on the fast ions particle flux. Shown here are 1st (red), 2nd (blue) and 5th (green) nearest neighbors and the result when all neighbor effects are taken into account (yellow)*

4.3 Benchmark Results

At this point two different models to simulate the heating of the plasma through collisions with fast alpha particle produced in fusion events were developed. To evaluate the effectiveness and correctness of these models they are compared with the results from the earlier mentioned Monte-Carlo code SPOT (Schneider *et al.* [9]). This comparison is done on the basis of energy deposition profiles to the electrons and the total (ions + electrons) energy deposition. Figure 4.2 shows these profiles for SPOT, the direct deposition model and the simple diffusion model at 3 different times during the ramp up. This is to ensure that different regimes of the important plasma profiles are tested.

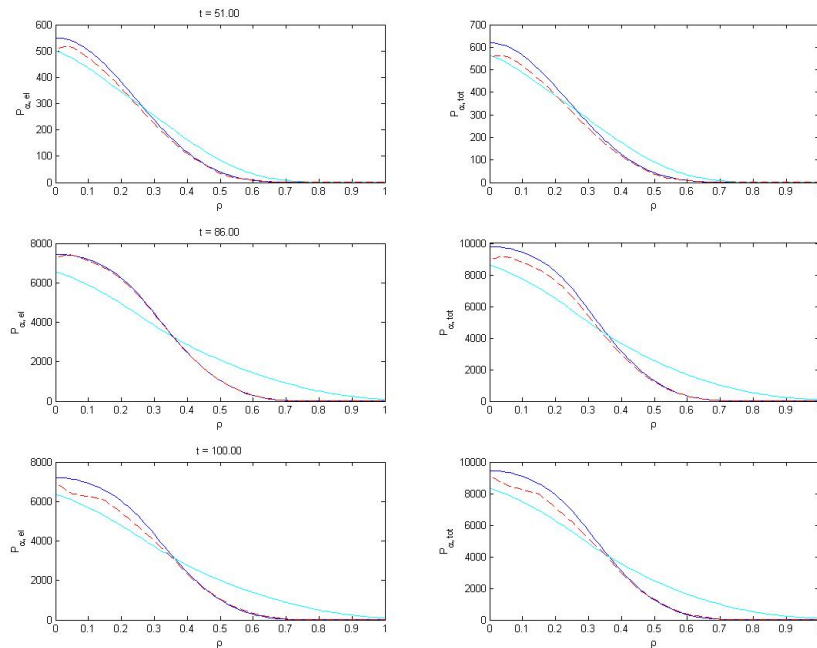


Figure 4.2: The energy deposition profiles at 3 different time steps for SPOT (red dashed line), direct model (blue line) and the simple diffusive model (cyan line)

It can be seen that the simple diffusion and the direct deposition model give very different profiles with a clear diffusion signature when these models are compared with respect to each other. However, neither models seem to be particularly better when compared to the SPOT deposition profile. It can be seen that the diffusion turning point (where the particle flux becomes zero),

as seen in the simple diffusion model (around $\rho \approx 0.4$) is found in the SPOT model results at a different points (around $\rho \approx 0.6$). It can also be seen that the fraction of energy deposited to the electrons and ions seems to be different in SPOT at certain times. This effect seems to have a larger impact than the diffusion. Therefore a correction factor in the calculation of the energy fraction to the electrons was added and tested. This can be seen in figure 4.3:

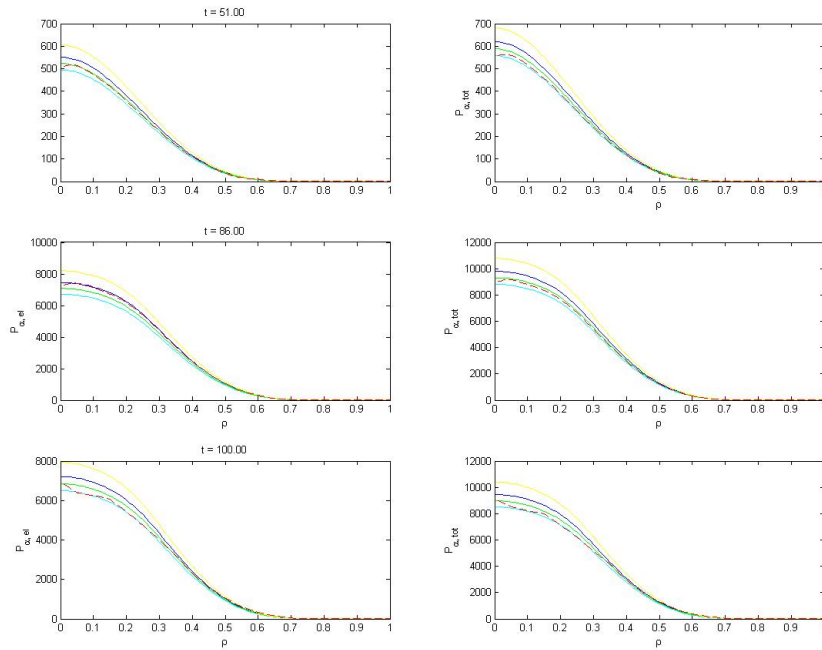


Figure 4.3: *The energy deposition for different adaption factors: 110% (yellow line), 100% (blue line), 95% (green line) and 90% (cyan line) of the original profile.*

It can be seen that this correction is more effective than the introduction of the diffusion and that the simple model now seems to effectively model the heating of the plasma through fusion events. To also quantitatively analyze the model the profile of the relative error between the simple model with an adaption factor and the SPOT model was determined. The result is shown in figure 4.4:

The fact that the correction factor is more effective than the diffusion model might be explained by the fact that there is a delay in the energy deposition but no large diffusion. This delay can be of the order of the evolution time step and is not modeled in RAPTOR. This difference can be slightly corrected

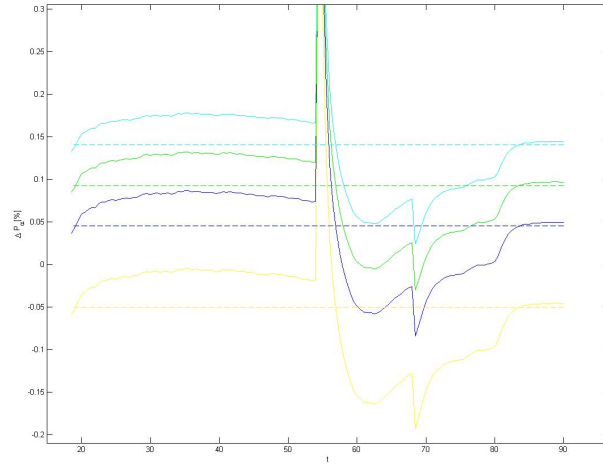


Figure 4.4: *Relative error for the different adaption factors, using same color coding as , dashed line shows average of relative error*

for a correction factor of 5% as seen in figure 4.4. The peaks in this figure are caused by the switch on of NBI heating, causing a temporal and localized increase of fusion events.

Chapter 5

Validating The RAPTOR Code For ITER

As mentioned before, two different simulation codes are used during the work described in this thesis. The RAPTOR code was previously validated for the TCV tokamak using the simulation code ASTRA [10]. The work done in this thesis is based on the international collaborated tokamak ITER which is going to be different from the TCV tokamak in many aspects. The simulation code RAPTOR should therefore be adapted to include all important physics relevant to the ITER tokamak and this update should also be validated. The following sections describe this validation.

5.1 Introduction

This validation is done using the fully integrated simulation code CRONOS, which was previously mentioned. There are several differences between these simulation codes. One of the main differences is that the geometry of the nested flux surfaces are computed, except for the boundary, in CRONOS whereas RAPTOR uses a prescribed equilibrium. This prescribed flux surface equilibrium is taken, for a particular scenario, from the CRONOS results and is kept fixed in all ITER simulations of that scenario in RAPTOR. This means that the 2D Grad-Shafranov equilibrium (eq 2.12) is not evaluated but presumed known. By also fixing the enclosed toroidal flux the spatial distribution of ρ_{tor} are fixed. This in turn has the consequence that the geometric quantities such as G_1 , G_2 and V' are fixed. This assumption was analyzed in the original development of RAPTOR [10]. It was found that differences in these profiles are limited and acceptable since a doubling of β resulted in a variation of G_2/J and V' of less than 10%. This difference

is one of the places where the physics in RAPTOR is different from the more extensive CRONOS. Another difference is the fact that the evolution of the ion-temperature is also pre-described. This can potentially have a great impact on the evolution of the discharge through heat exchange with the electrons. Other differences will become apparent during the rest of this comparison.

Within the CRONOS code, it is possible to choose which profiles need to be evaluated and which can be prescribed by predetermined profiles. During the validation process we started by creating a baseline in which the equations evaluated in CRONOS match those considered in RAPTOR as much as possible. The most important nonlinear coupling between plasma parameters during the evolution of tokamak profiles originate from the electron temperature and q-profile. The particle density profiles are much harder to model and more uncertain. Also the coupling of the profile with other plasma profiles is quit weak. Therefore the density profiles are usually not modeled but prescribed. The same holds for the ion densities with the exception that the radiation from plasma ions can be important. The ion and electron density profiles are also correlated because of quasi-neutrality. Therefore these can be approximated by typical profiles and be pre-described effectively. The main evolution equations that are therefore considered in RAPTOR are the energy transport equation (eq 2.48) and the poloidal flux diffusion equation (eq 2.45). These equations hold several quantities which are addressed differently in both simulation codes. These differences were analyzed in detail during the work in this thesis and .

In the case of the poloidal flux diffusion equation the most important quantity is the parallel conductivity $\sigma_{\parallel}(T_e)$, the bootstrap current and the geometric factors mentioned earlier. Since the differences arising from geometric factors were analyzed previously they are not considered here. The geometry in the comparison runs are chosen such that they resemble the RAPTOR equilibrium in the flat top of the CRONOS simulations. The conductivity and the bootstrap current were considered during this validation and will be addressed in the following sections.

In the case of the energy transport equation the heat and particle fluxes are important as well as the sources and sinks considered. In the case of the heat flux there are two important influences which need to be validated. Firstly, the heat fluxes that are considered in RAPTOR is just the heat flux due to electrons whereas CRONOS can also consider the ion temperature (and its evolution/heat flux) as well as the heat exchange between electrons

Physics	RAPTOR	CRONOS
Heat flux (q)	χ_{BgB}, T_e	χ_{BgB}, T_e, T_i
Particle flux (Γ)	NB	n_e
P_{source}	$P_\alpha, P_{brem}, P_\Omega, P_{ei}$	$P_\alpha, P_{brem}, P_\Omega, P_{ei}, P_{rad}$

Table 5.1: Schematic representation of simulation differences between RAPTOR and CRONOS

and ions (when at different temperatures). Secondly the transport model used to model the energy transport can be different. To minimize this difference, the transport model used to evaluate the energy transport was chosen identical. The transport model used was taken from literature and developed and tested for RAPTOR during the work on this thesis. As mentioned before, another difference in the energy transport is the particle flux which can be evaluated in CRONOS but is not available in RAPTOR. Lastly the energy sources and sinks considered can, in principle, be different or their influence can be evaluated differently. These differences were also considered and all the differences described above will be discussed in more detail in the following sections. The differences are given schematically in Table 5.1:

5.2 Baseline Comparison

To start the comparison both models were stripped down to contain not much more than the energy transport and parallel conductivity. This was used as a baseline from which further comparison is done. Additional ingredients to the full set of differential equations were added and their influence analyzed. In this baseline comparison two major differences (besides the differences already mentioned) and an important equivalence were identified. The important equivalence is the energy transport model. We use a model described in literature [11] that is also implemented and tested in CRONOS. This model was also developed and tested for RAPTOR and were found to be equivalent. However, major difference were found in the calculation of the neoclassical term in the conductivity. Two different models are used for the calculation of this correction term. Where CRONOS uses the NCLASS code described in [12], RAPTOR uses the treatment based on Sauter(SA) *et al.* [13]. Due to the coupling between the two differential equations the difference in σ has effect on both the evolution of ψ and T_e . A last major difference is the numerics used in both codes. Where CRONOS uses RAPTOR on the other hand uses spline functions and a general Crank-Nicholson discretization to arrive at Ordinary Differential Equations (ODE's) which are

solved using the previously mentioned Cholesky decomposition. The latter differences can not be analyzed easily since they are inherent to the codes used and have to be accepted as intrinsic features. For more information about the numerics of both codes the reader is directed to [10] (RAPTOR) and [14] (CRONOS). The (re-)development of the treatment of the transport model and neoclassical conductivity is discussed in the following sections.

5.2.1 Transport Model: Bohm-GyroBohm

Over the years, many different energy transport models have been developed based on very different physical or experimental arguments. We have chosen to use a transport model that has already been implemented in CRONOS and has proven its value in experiments from the ITER database, JET and Tore-Supra discharges [11] and is widespread in the fusion community. This model is based on transport on two length scales. The first is the scale of the gyration around the magnetic field lines (gyro-Bohm, given in eq (B.7)) and the second on the length scale of the tokamaks' minor radius (Bohm, given in eq (B.6)). By mixing the dependence on these two length scales, the authors of the article show that the transport model can predict the transport of a wide range of machine parameters. This transport model has shown to hold the correct scaling properties that allow it to work in different machines and represents a possible general transport model. A more elaborate derivation and description of the Bohm-GyroBohm transport model is given in the Appendix.

5.2.2 Conductivity

The calculation of the transport parallel to the magnetic field lines is believed to be well described by the neoclassical theory. The neoclassical conductivity and the bootstrap current are used in a wide range of tokamak simulations. To calculate the neoclassical conductivity for a wide range of plasma parameters one needs to solve a set of multi-species fluid equations using multiple odd velocity moments of the Fokker-Planck equation. The NCLASS solves this set of equations but is inconvenient for use in rapid codes. It is also known to not use the full collision operator which can, in some cases, lead to errors as big as 20% according to [13]. The calculation of the neoclassical conductivity as described in [13] is based on the code CQLP and CQL3D which uses the linearized operator to solve the Fokker-Planck equation. The SA neoclassical conductivity and bootstrap current are based on a polynomial fit of the geometry (fraction of trapped particles) and collisionality (ν_*), which is in turn based on plasma parameters such as density and tempera-

ture) on the results from the CQLP and CQL3D codes. The main question in this section is how to compare results from CRONOS and RAPTOR using these different neoclassical corrections. An additional question that needs to be asked is which of these approaches gives the best results for plasma parameters in the ITER regime. This is however a very complex question and the details of these different models are beyond the scope of this thesis. Therefore a brief argumentation from the authors of the papers/models in question is given:

-

Next, the difference in outcome of these models and its impact on the evolution of the profiles is discussed. During the validation procedure it was found that results for the conductivity were different by approximately 20%, the difference is illustrated for a typical ITER scenario profile in figure 5.1:

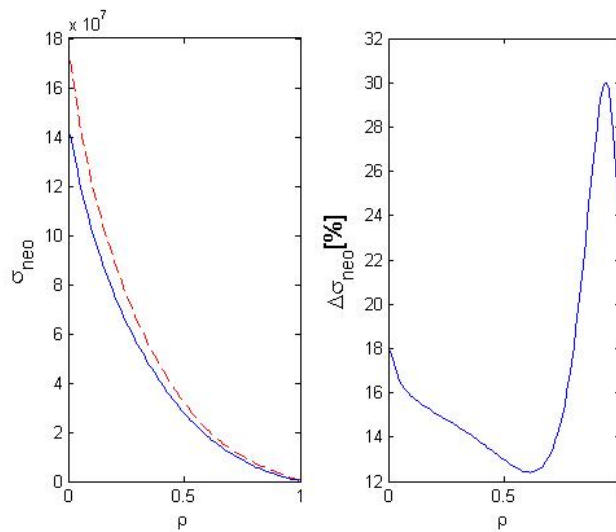


Figure 5.1: Left figure shows the neoclassical resistivity for a typical profile calculated using NCLASS (red dashed) and Sauter (blue). Right figure shows the relative difference between the profiles

This difference in neoclassical conductivity causes a difference in the evolution of both the ψ and T_e profiles. Which in turn causes a difference in the conductivity through its temperature dependence, causing a feedback effect. Determining its effect on the complete evolution of the profiles is therefore complex and needs to be dealt with cautiously. Another problem in analyzing this error is the fact that the numerical computation is different and that

these errors occur simultaneously.

The conductivity is usually calculated by using the following form [1]:

$$\sigma_{\parallel} = \sigma_{neo}\sigma_{spitzer}. \quad (5.1)$$

where $\sigma_{spitzer}$ is the standard classical conductivity and the neoclassical conductivity is the part that is different for the models used. Originally, RAPTOR used a typical profile (for a given tokamak) to calculate the neoclassical conductivity and used this profile throughout the simulation. This would normally be acceptable because the neoclassical correction does not vary much. It was found that a typical ITER-like simulation goes through such a wide range of plasma parameters that this approach was no longer acceptable. Therefore the model had to be modified accordingly such that it accommodates the ITER plasma regime.

5.2.3 Evolution Comparison

Now that the differences in the models describing the transport in the evolution equations have been discussed it is time to look at the comparison between the RAPTOR and the CRONOS results. The baseline results consists, as already mentioned, of nothing more than ohmic heating and current drive and, in the case of RAPTOR, a pre-described magnetic equilibrium. This gives a good insight into the differences in the profiles from the basis of the two codes. In this comparison the profiles that we are looking at are representative for the difference in any of the other profiles. These are the electron temperature, ι , the plasma current I_p and the plasma loop voltage U_{pl} . The following figure 5.2 shows, for each profile, the results from RAPTOR and CRONOS at 4 different times. It can be seen that the codes seem to be in good agreement from visual inspection:

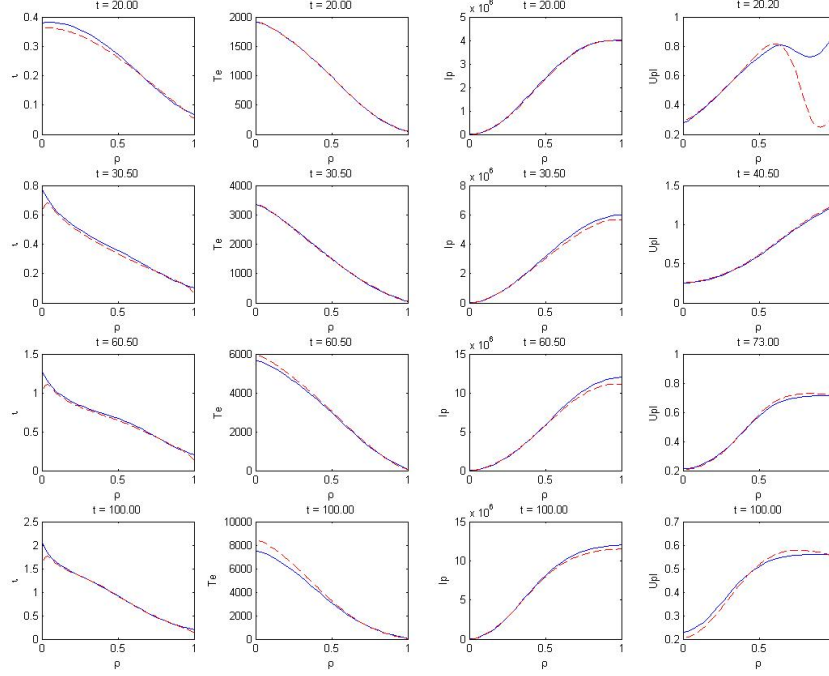


Figure 5.2: The comparison of 4 different profiles at 4 different time steps. Solid blue line represents the results from RAPTOR and the dashed red line represents CRONOS results.

The profiles seem to be in good agreement but to be more certain a quantitative analysis of the results was also done. This quantitative analysis was performed by calculating the relative difference of the RAPTOR profile with respect to the CRONOS profile. The mean of this error (Eq 5.2) and its deviation say something about how good the profiles match each other. It is insightful to look at the time dependence of the ρ -averaged difference:

$$\Delta O(t) = \frac{\sum_{\rho} \frac{O_{rp}(\rho,t) - O_{cr}(\rho,t)}{O_{cr}(\rho,t)}}{N_{\rho}} \quad (5.2)$$

Throughout this comparison, the profiles will be quantitatively compared using this method. For clarity, these results will be compared to (at least) this baseline result. The time dependent result for the baseline comparison is shown in figure 5.3:

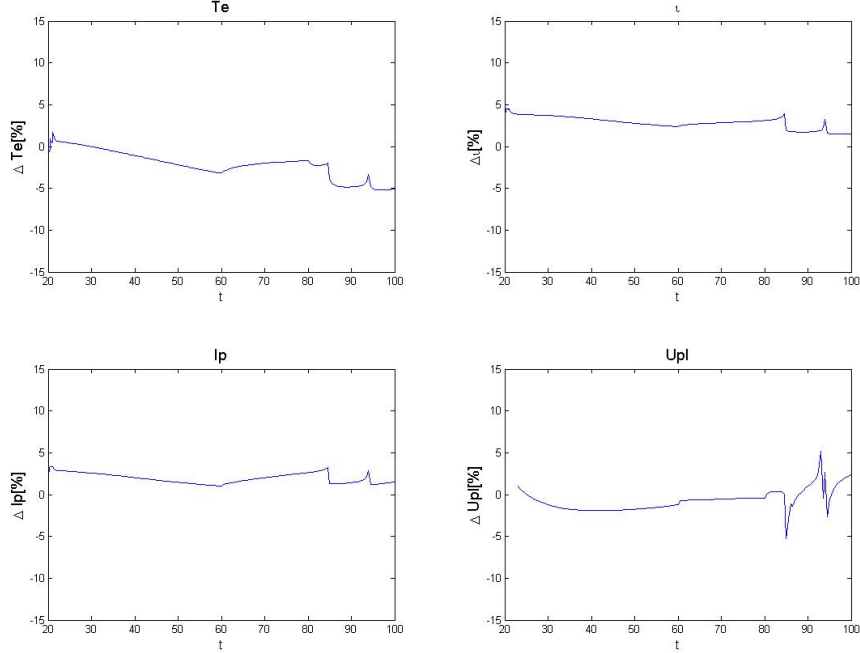


Figure 5.3: *The results from ρ -averaged relative difference between RAPTOR and CRONOS.*

These relative differences show that the profile difference between RAPTOR and CRONOS are approximately 5% for the ohmic heating scenario, where a correction factor of 15% was used to alleviate the difference in the treatment of the neoclassical conductivity. A distinctive kink can be observed at $t = 60$ s but this can be associated to turning point for the ohmic driven current going from uniformly increasing to flat-top. Another distinctive feature of these profiles is the discontinuity at approximately $t = 84$ s. This is caused by numerical instabilities in CRONOS for which the reasons are unclear and which could not be resolved. Since the flat-top was reached far before this point we have been able to establish, with good confidence, that the behavior of these difference profiles behave normally in the flat-top and that no major differences arise.

The differences that do arise are, as mentioned, in the order of 5% and can only be caused by differences in the neoclassical corrections of the conductivity and numerical treatment of the evolution. The latter can not be influenced and thus be analyzed but the neoclassical correction can be adapted

by a constant multiplication factor. The effect of tuning this factor is shown in figure 5.4:

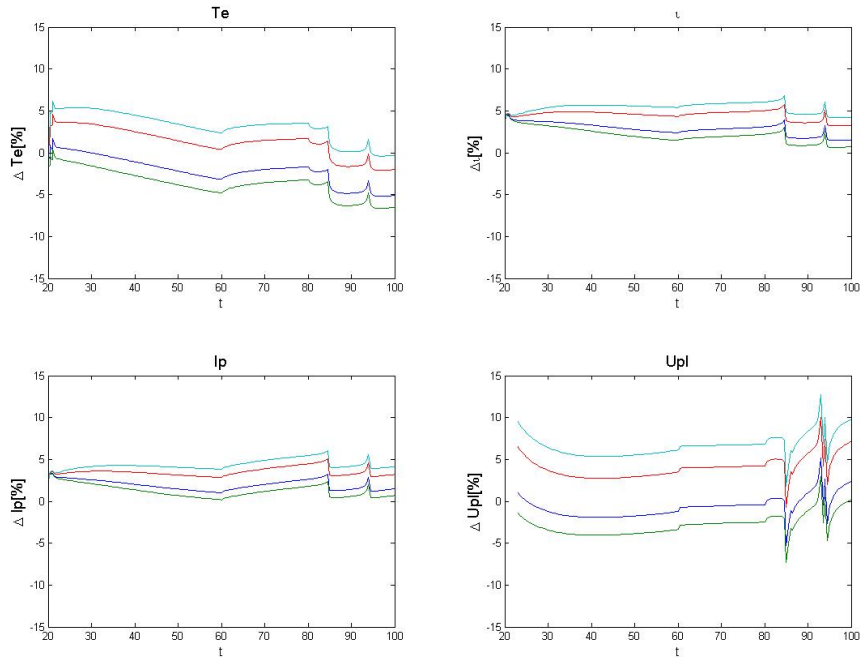


Figure 5.4: A comparison of the relative differences results for different σ_{neo} adaption factors. Blue: original 15%, Green: 20%, Red: 5% and Cyan: 0%

5.3 Poloidal Flux Diffusion

The poloidal flux diffusion equation (Eq 2.45) holds, apart from the conductivity another two ingredients that might alter the evolution during the simulation besides the earlier mentioned G_1 , G_2 and J . These are the sources for current density from auxiliary sources and the bootstrap current. The auxiliary current is normally always combined with an energy source in the energy transport equation. Here we bench marked the auxiliary current and heat sources also in combined fashion. This section provides an insight into the bench marked bootstrap current only but the next section will treat the auxiliary energy source and corresponding current drive.

5.3.1 Bootstrap Current

As for the conductivity, the bootstrap current is also a neoclassically calculated quantity with turbulence induced variations. It is therefore subject to the same form of calculations as the conductivity and the same difference arises. Also the bootstrap current in RAPTOR is calculated using the fit found in Sauter [13] whereas CRONOS uses the NCLASS routine to calculate the neoclassical bootstrap current. This means that the different calculation procedures yield differences in evolution which can not be excluded. Also, RAPTOR also uses a typical profile to calculate the bootstrap current here but since the bootstrap current in L-mode does not contribute greatly this will not affect results significantly. However, when the H-mode regime is entered this is a different story and the bootstrap needs adaptation. The comparison was done in L-mode and the results were equivalent to the baseline comparison. No additional error was introduced by the bootstrap current, mainly because its contribution is negligible in L-mode.

H-mode adaption for Bootstrap current. To be written.

5.4 Energy Transport Equation

The energy transport equation does not differ from CRONOS apart from the already mentioned transport model. There are the energy sources and sinks that are to be considered and their implementation that can differ between RAPTOR and CRONOS. There are inherent power sinks and sources to the plasma physics such as radiation losses, particle flux and heat exchange between plasma species. On the other hand there are also multiple auxiliary power sources that can be applied to heat the plasma. These different methods were already discussed in an earlier stage. The difference between RAPTOR and CRONOS in the way they treat these external power sources, along with the intrinsic power sources and sinks, will be discussed in coming sections.

5.4.1 Particle Flux

In Eq 2.48 there are two different flux terms: the heat flux (eq 2.50) and the particle flux (eq 2.49). In RAPTOR, only the first term is taken into consideration because this will be the main contribution. However, it can be shown from the profiles used in CRONOS that the energy flux due to the particle

flux can be a contribution between -6% & 8% but it averages out over the profile then not much more than 4% of the energy transport. That should mean that the energy transport equation is not affected much by neglecting the particle flux. In figure 5.5 the relative contribution of the particle flux to the time-derivative of the energy content of the plasma is considered with respect to particle and heat flux. In this picture, the divergences have been cut out to give a clearer picture:

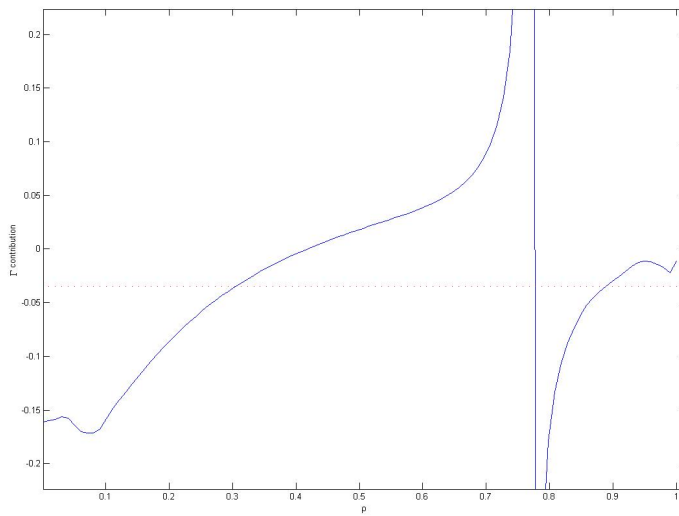


Figure 5.5: *The contribution of particle flux relative to total energy flux for a typical ITER profile. The red dotted line shows the average over the profile of -3.5% .*

To check the influence of this affect on the total evolution of the profile during the ramp-up phase we have set the same conditions as in the baseline test but now added the evolution and contribution of the particle flux. In the following figure (5.6) the results of this simulation were added to the original results from figure 5.3:

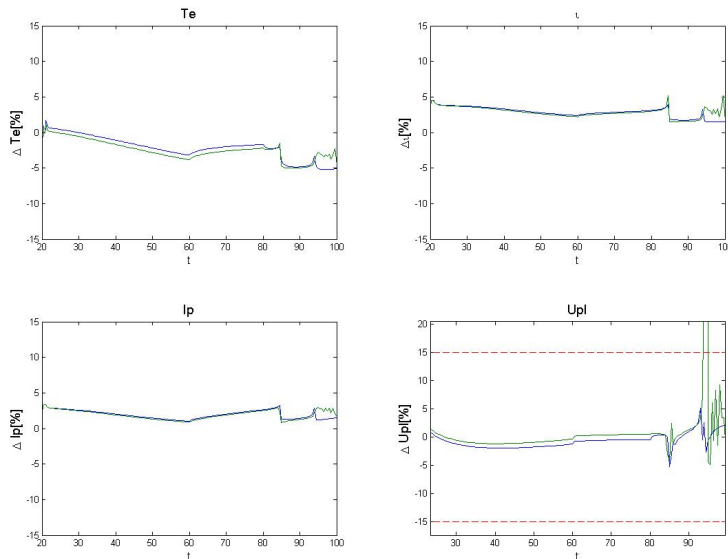


Figure 5.6: The result when introducing the particle flux power loss (green line), compared to the baseline results (blue line).

It can be seen that the effect of adding the particle flux to the energy transport equation affects the T_e and U_{pl} profiles with respect to the baseline run but only slightly and this effect is negligible.

5.4.2 Ion-Electron Heat Exchange

The ion-electron heat exchange occurs when different species in the plasma have different temperatures. In a typical ITER discharge, the ion temperature is expected to be lower than the electron temperature by approximately 10 – 20%. Due to this temperature difference there will be a power loss for the electrons and a power gain for the ions. In RAPTOR, the ion temperature is not simulated but the power loss to the electrons can be of significant magnitude depending on the ion temperature. The heat loss for the electrons due to this process is given in equation (2.51)

The fact that the ion temperature is not simulated in RAPTOR makes it difficult to correctly simulate this power loss. The choice was made to take the ion temperature as a fraction of the electron temperature. Thus assuming that the profile has the same shape but a space-independent constant multiplication factor that induces this loss process. When the factor changes greatly over time or space due to, for instance, only heating to the elec-

trons/ions this factor can be adapted.

The effect on the difference of implementation was first evaluated using a CRONOS run which had predetermined ion temperature and density profiles such that the ion temperature was 90% of the original electron temperature. During this new run, only the electron temperature was simulated to see what the effect would have been on the electron temperature and no additional heating was used. The results, compared to the baseline comparison, can be seen in figure 5.7:

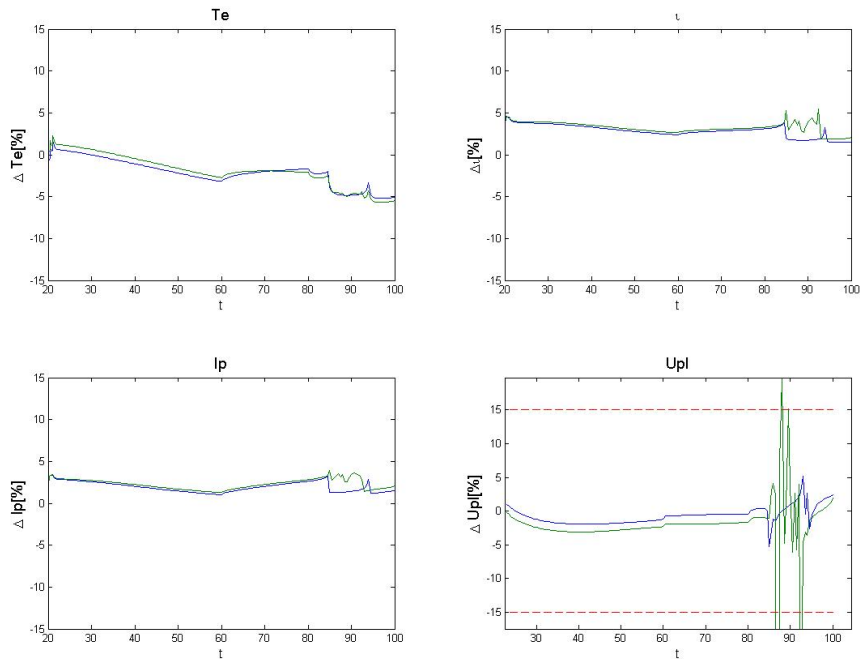


Figure 5.7: The result when introducing the heat exchange power loss (green line), compared to the baseline results (blue line).

Next is a comparison of the same RAPTOR run with a CRONOS run in which the ion temperature was actually simulated. It was found that the ratio changed a little over time. However it did significantly changed over the plasma radius. Therefore the correction factor was made linearly dependent on the plasma radius. This new ratio was 90% in the center and 100% at the edge. A new adaption coefficient was chosen such that it would match the ratio between the ion and electron temperature of the CRONOS run. The results of this comparison are shown in figure 5.8:

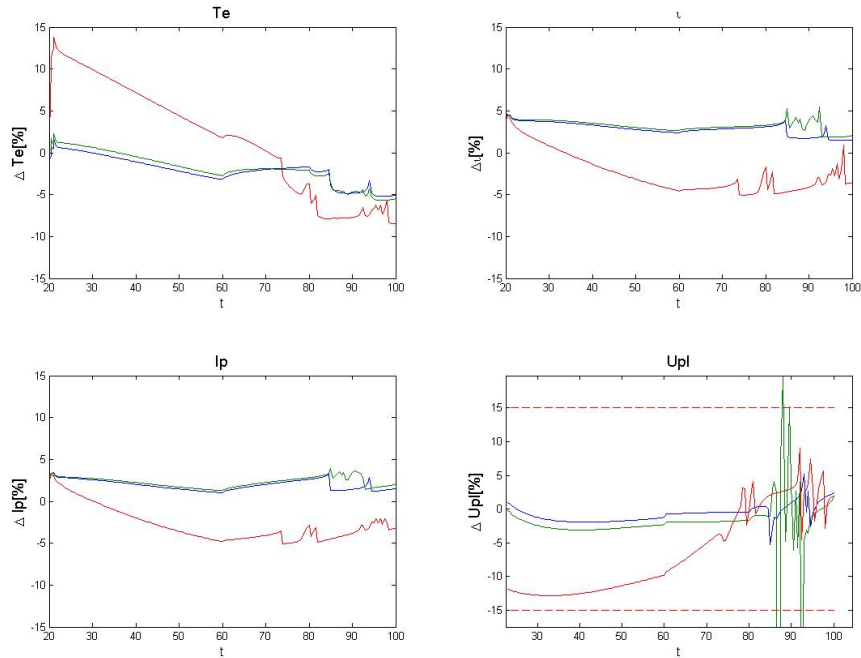


Figure 5.8: *The result when introducing the heat exchange power loss (red line), now for a simulated ion temperature, compared to the baseline results (blue line).*

5.4.3 Radiation Losses

Multiple types of radiation play a role in tokamak plasma physics but only 2 can cause significant losses. These are the bremsstrahlung, caused by acceleration of a charged particle, and the line radiation which is most significant for the impurity ions. The power loss due to the bremsstrahlung is given in Eq (2.52) and is affected by the impurities but only globally through Z_{eff} . The line radiation however is affected by the impurities in a more complex manner. Information about each impurities' temperature, density and ionization is needed to be able to construct a full power loss profile. Such an extensive calculation is beyond the scope of the RAPTOR code and would significantly slow down the computation, making it less useful for the optimization. The bremsstrahlung loss was implemented in RAPTOR as part of this thesis using expression (2.52). The effect on the evolution of adding this radiation can be seen in figure 5.9:

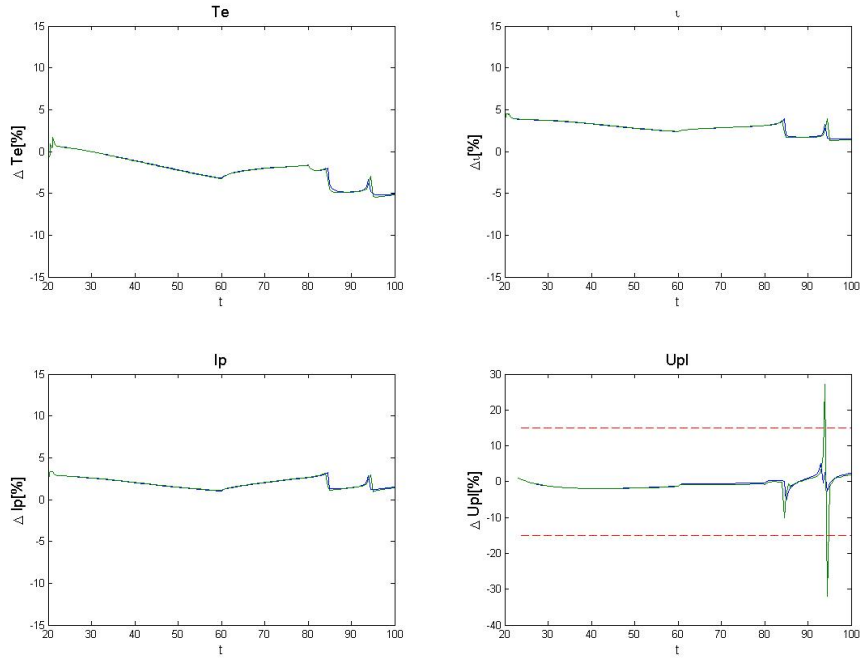


Figure 5.9: The result when introducing the bremsstrahlung power loss (green line), compared to the baseline results (blue line).

It can be seen that the difference is minimal when introducing the bremsstrahlung radiation. It can therefore be assumed that the bremsstrahlung radiation model in RAPTOR is equivalent to CRONOS. As an additional test we also compare the power loss profiles. The results for 3 different time steps are given in figure 5.10. The remaining difference in these profiles stem from differences in the evolution of the electron temperature.

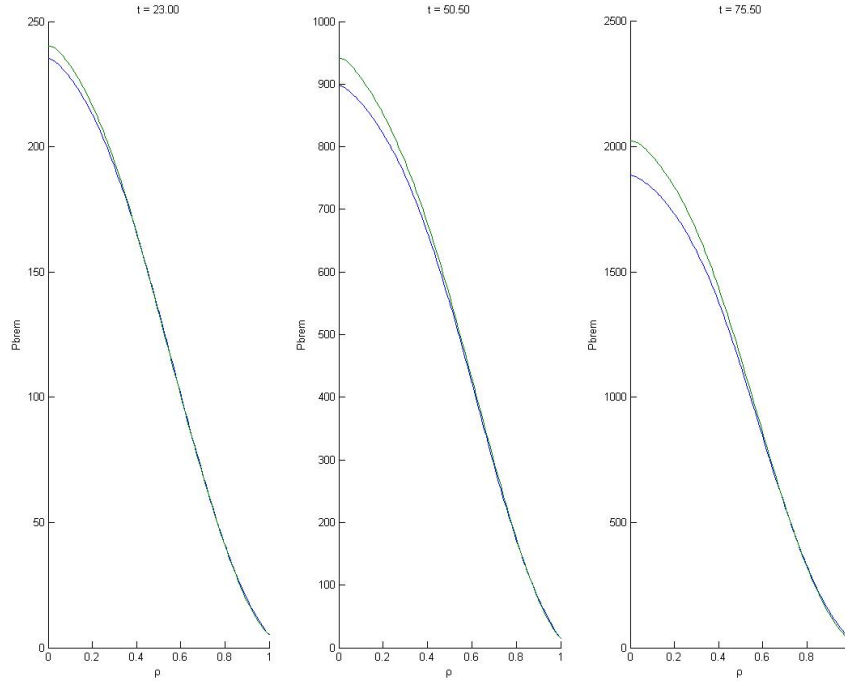


Figure 5.10: Power loss profiles due to bremsstrahlung at 3 different time steps for RAPTOR (blue) and CRONOS (green)

The line radiation is a more complex process. Although it is hard to model exactly, it does have a significant affect on the discharge. Therefore it is necessary that a loss term is modeled that mimics the physics of line radiation. This is done in RAPTOR by adding a loss term that is proportional to the ohmic power to the electrons. In addition to this, the peaks of line radiation from the largest contributor (Argon) are modeled by gaussian peaks. These peaks are centered around certain temperatures at which the contributing ions radiation is peaked. The multiplicative factors to these peaks are proportional to the ion density and chosen constant factor such that the peaks match the results from CRONOS as good as possible. The results from this procedure can be seen in figure 5.11:

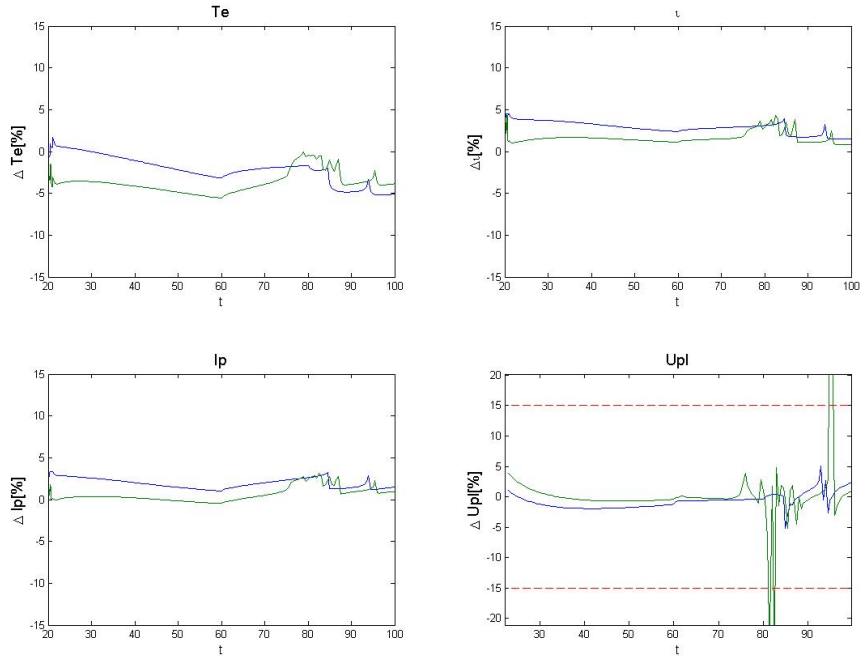


Figure 5.11: The result when introducing the line radiation power loss (green line), compared to the baseline results (blue line).

5.4.4 Auxiliary Heating

The auxiliary heating plays a crucial role in the heating of the plasma discharge. It is therefore vital to properly model the physics of the auxiliary heating to be able to do the optimization. As a first comparison we will introduce to RAPTOR the heating and current drive profiles due to electron cyclotron resonant heating (ECRH/ECCD). By using the profiles given in CRONOS the difference due to the way of calculating these profiles is eliminated and the effect of heating/current drive on the evolution is isolated. The next step is to let RAPTOR calculate the deposition profiles itself but still using similar models (gaussian distribution). The final step in this comparison will be between the RAPTOR model (Gaussian distribution) and the more complex model used in CRONOS. Another means of auxiliary heating was also developed for RAPTOR, parallel to the work done in this thesis. This model was bench marked by P. Geelen [] in his thesis. The difference between this model and the model used in CRONOS is also analyzed here.

For the electron cyclotron resonant heating, we have chosen a rather er-

ratio profile to make any differences apparent. The heating profile used is shown in figure 5.12:

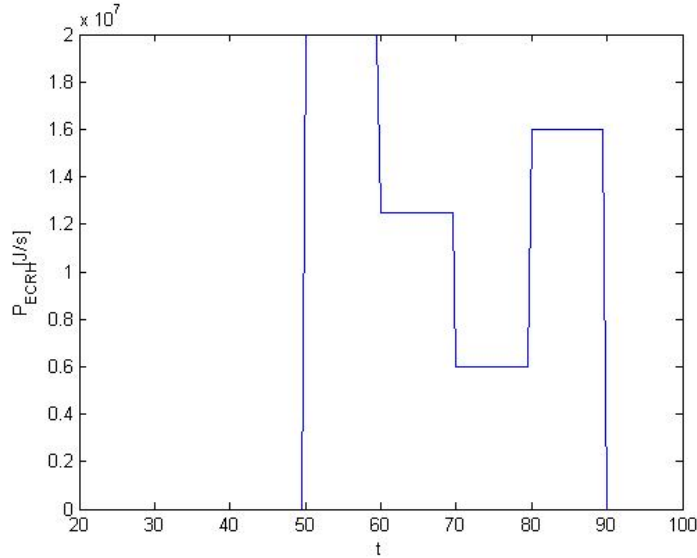


Figure 5.12: *Energy deposition over time*

Firstly, the most general case, where the deposition profiles from CRONOS are used, will be analyzed. The results can be seen in figure 5.13. This shows the results with heating compared to the baseline results. It can be seen that the heating causes an increase in the errors, especially in U_{pl} , but that the errors are still within the acceptable range. Also it can be seen that the errors increase when the input power is changed.

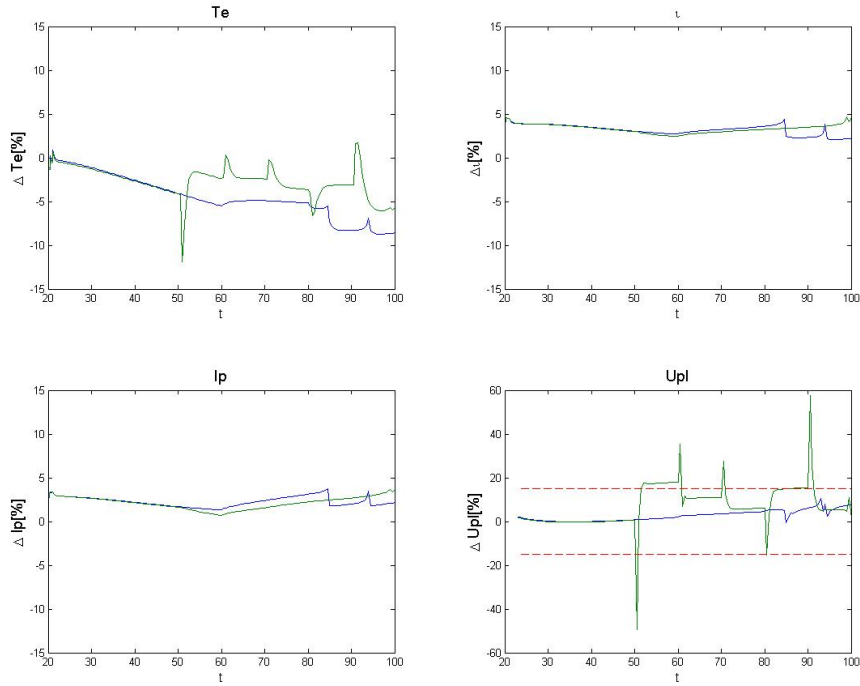


Figure 5.13: The results of adding a heat source to the evolution (green) compared to the baseline (blue)

Next up is the comparison between similar models which give a gaussian deposition profile at a given radius and with a given width. In principle, both models should give the same (or at least similar) deposition profiles. Since the difference for identical profiles is known from the previous result this will show a difference in the calculation of the deposition model. The following figures show: (5.14) the difference between the calculated deposition profiles and (5.15) the impact on the evolution of the profiles. It can be seen that the deposition profiles for these models are very similar and that the error caused is therefore also quite similar. There is however a small difference in the error of the temperature profile and in the plasma loop voltage.

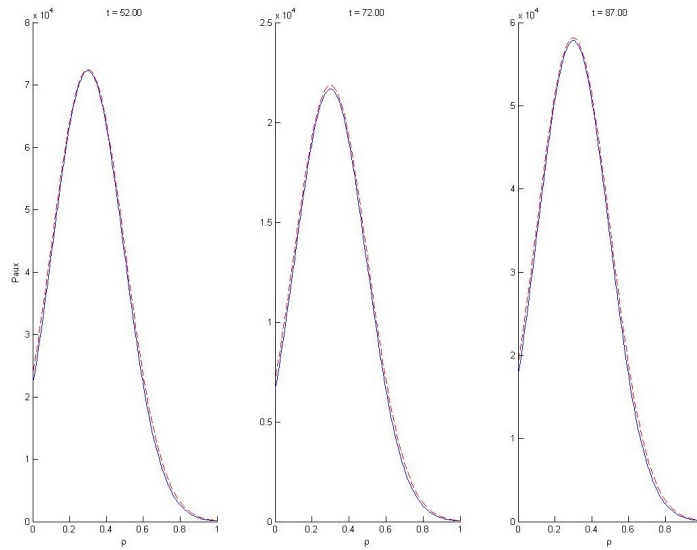


Figure 5.14: Gaussian deposition profiles calculated in CRONOS (red dashed line) and RAPTOR (blue line)

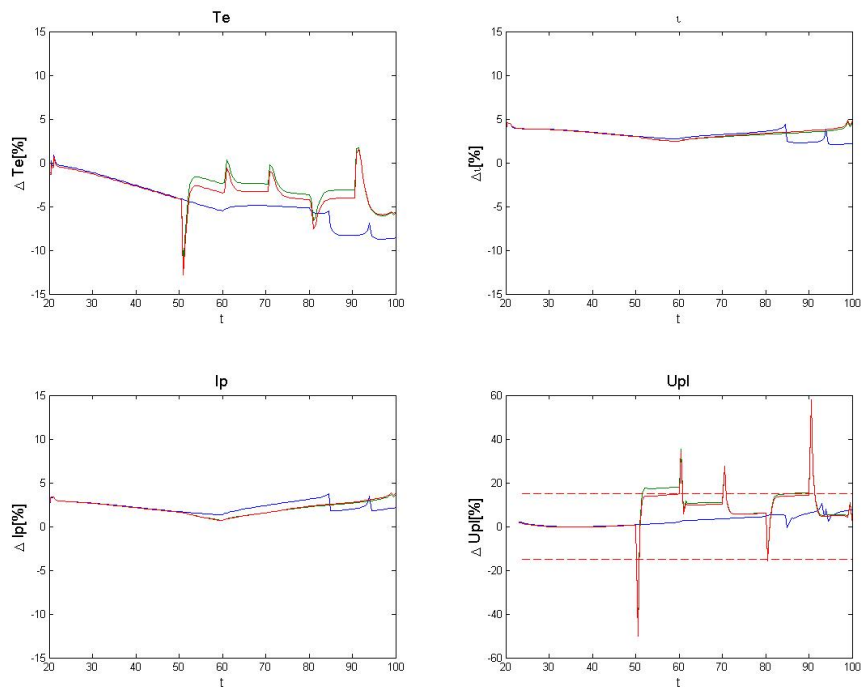


Figure 5.15: Similar figure as 5.13 but now using RAPTOR ECCD/ECRH heating model (red line)

The CRONOS code also has a model which is based on experimental parameter such as inclination and frequency of EM waves. This was it is more realistic model for the power deposition but also more difficult to control and model. This is however a realistic test for the way this heating his controlled in the future by experimentalists. It is therefore useful to identify the differences in the calculated deposition profile by this model and RAPTOR's model. The difference in these profiles can be seen in figure ???. Also the impact on the evolution gives, again, valuable insight, this is shown in figure ???:

It can be seen that and bla bla bla

Now the same results are depicted for the comparison with ECCD (electron cyclotron current drive) using the same procedure only now show the results of the first two steps in a single picture in figure 5.17. First the comparison between the CRONOS deposition profiles and the RAPTOR model is shown in figure 5.16:

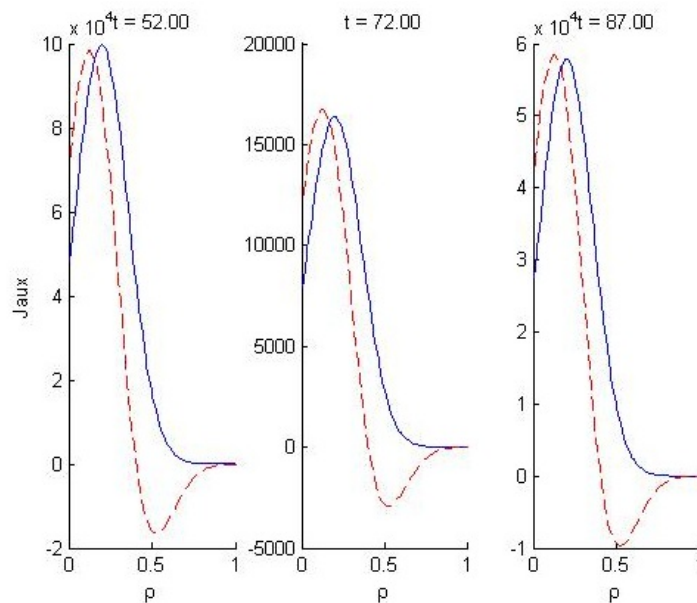


Figure 5.16: Current drive deposition profiles for the gaussian RAPTOR model (blue lines) and the gaussian CRONOS model (red dashed lines).

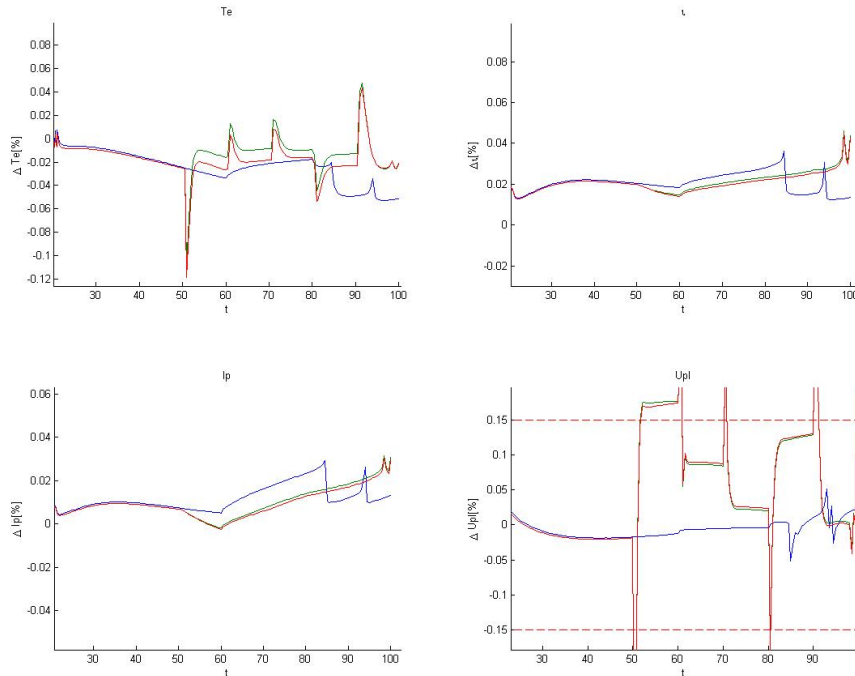


Figure 5.17

It can be seen that the results are very similar to that of the ECRH heating. The deposition profiles for the current drive are, however, not as similar as the power deposition is. The fact that this does not seem to have a large effect on the evolution suggests that the current drive is not of great influence. It was found that the current drive from the ECCD is not more than a few percent of the total current density. In figure ?? the comparison is shown for the complex ECCD model from CRONOS with respect to the gaussian RAPTOR model:

Another important heating scheme is the injection of energetic neutral beams. As mentioned before, this was implemented and tested for RAPTOR by P. Geelen. The comparison between this model and the most sophisticated model in CRONOS is however done here. The results can again be seen in the same format in figure ??:

It can be seen that bla bla bla and bla bla bla

5.4.5 Alpha Particle Heating

As described in Chapter 4, important work in this thesis was to implement the process of alpha particle heating of the plasma as a result of the fusion process. The implementation and bench marking of the model used in RAPTOR was described in Chapter 4. It was however not yet shown how this interaction influences the evolution of the T_e and ψ profiles with respect to the CRONOS evolution. This section will show the results from the comparison for the alpha particle heating.

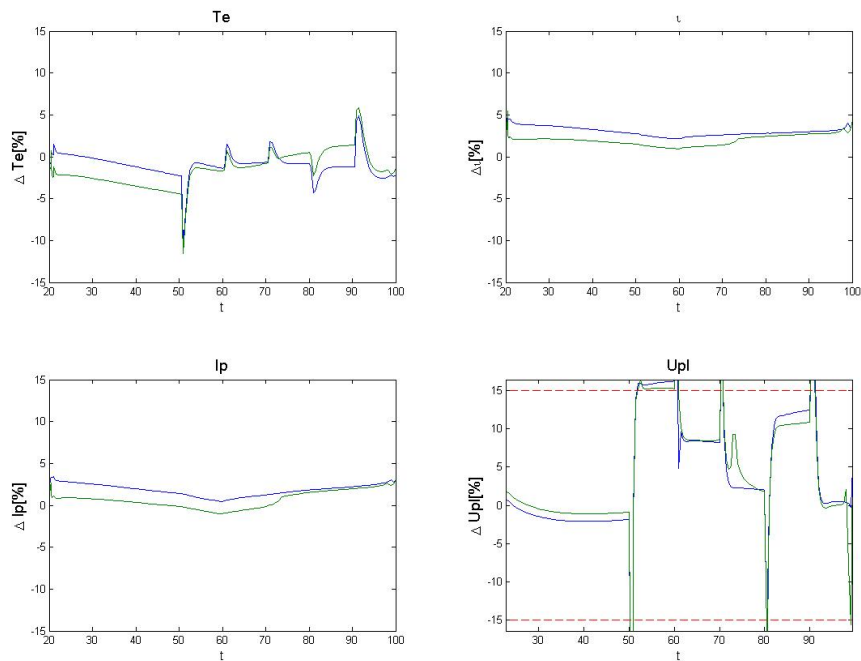


Figure 5.18: This figure shows the influence of introducing the fusion power (green line) with respect to the simulation with only heating (blue line).

Figure 5.18 shows the results of including the alpha particle heating to the electrons in the model. This is done with respect to the ECCD heated evolution. This is because the alpha particle production needs considerable pressure to be significant in the heating of the plasma. It can be seen that introducing this effect has only little effect on the evolution compared to ECCD heated evolution.

5.5 Full Transport Model Comparison

At this point it is time to test the RAPTOR model with all the necessary ingredients combined. This means that all different interactions are taken into account in the following comparison. However, we do want to look at the effect of the electron-ion heat exchange separately because of the pre-described ion temperature and density profiles in RAPTOR.

5.5.1 Electron-Ion Heat Exchange

As previously for the ohmic heating case, we now investigate the impact of also introducing the electron-ion heat exchange to the evolution of the full model. The same procedure of comparison for this plasma interaction is used as before: First use an ion temperature of 90% of the original electron temperature in CRONOS and secondly, let CRONOS evaluate the ion temperature and adapt the RAPTOR coefficient to match the average temperature difference.

Chapter 6

Optimization

The previous chapter shows the verification of the validity of the RAPTOR simulation code with respect to the CRONOS simulation code. We are now in the position to start the numerical optimization procedure for ITER simulations. In this optimization procedure we will use the sequential quadratic programming (SQP) method to optimize the trajectories of the plasma current I_p and the ECCD/ECRH sources. At this point the NBI power is not yet optimized. This is mainly because the power, in ITER, can not be set at continuously distributed powers but only at full or half power. This complicates the optimization of this actuator and given the limited freedom for this actuator it was decided that we fix the NBI trajectory for now. The actuators are parameterized by a piecewise linear function. The discretization of this function is determined by the user by choosing a set of time grid points. The choice on the number of time grid points corresponds to a choice in number of optimization parameters. These parameters are then optimized in an optimization program using SQP. The resulting trajectories will also be simulated using the CRONOS simulation code to check their validity. Additionally, the sensitivity of the trajectories with respect to changes in the model will be investigated.

6.1 Optimization set-up

In order to do the optimization the resulting profiles from simulations have to be translated to a scalar which reflects their performance with respect to some requirements, introduced in Chapter 3 as the cost function. This cost function can be motivated by many different aspects of tokamak operation such as: instabilities or turbulence suppression, coil limits and efficiency. Additionally, some areas of the parameter space might be unfavorable or

inaccessible due to limits on, for instance, the coils or plasma stability. These parts of the parameter space can be avoided by imposing constraints on the actuators or on the plasma profiles. The set of possible cost functions and constraint together with their motivations are discussed in the following sections.

6.1.1 Cost-Function

In the original optimization work in the thesis by F. Felici [10] several different cost functions were motivated and tested. All these cost functions are dependent only on the final state of the plasma, $x(t_f)$, and its derivative $\dot{x}(t_f)$. Depending on the problem one is trying to solve, different contributions to the cost function can be taken into account. The cost functions are generally defined by a continuous norm function given by:

$$\|f(a, b)\|_W^2 \equiv \int_a^b [W(\rho')f(\rho')]^2 d\rho' \quad (6.1)$$

where $W(\rho)$ is a weight function and the function $f(\rho)$ describes the penalization with respect to the desired profile characteristic. As mentioned previously, several cost functions were defined in the original work. Their form and motivation were:

- The first cost function choose to penalize any deviations from a desired final rotational transform profile ι . It was defined as:

$$J_\iota = \|\iota(t_f) - \iota_{ref}\|_{W_\iota}^2 \quad (6.2)$$

- To reach a steady-state scenario, in which the operation could in theory be continuous instead of pulsed, the final loop voltage profile would have to be zero. Therefore any contribution to the loop voltage profile is penalized in this cost function:

$$J_{U_{pl}} = \|U_{pl}(t_f)\|_{W_{U_{pl}}}^2 = \|\dot{\psi}(t_f)\|_{W_{U_{pl}}}^2 \quad (6.3)$$

which is only dependent on the time derivative of the plasma state at the final time

- After the initial ramp-up phase, it is desired that a stationary state is reached. The degree of stationarity is determined by looking at the ρ derivative of the loop voltage profile. The cost-function is therefore given by:

$$J_{ss} = \left\| \left. \frac{\partial U_{pl}}{\partial \rho} \right|_{t_f} \right\|_{W_{ss}}^2 \quad (6.4)$$

- In order to maximize the duration of the flat-top operation phase one has to minimize the ohmic flux consumption. Therefore the consumption of ohmic flux is penalized by:

$$J_{\psi_{OH}} = \Delta\psi_{OH}^2 = [\psi_{OH}(t_f) - \psi_{OH}(t_0)]^2 \quad (6.5)$$

Which has a slightly different form than the general cost function form.

- It can also be desirable to penalize any deviations from a desired final electron temperature profile:

$$J_{T_e} = ||T_e(t_f) - T_{e,ref}||_{W_{T_e}}^2 \quad (6.6)$$

During the work on this thesis it was found that two additional cost functions can be preferable to the ones originally used. The first one is a different approach to penalizing the distance from a stationary state. This was needed because the original cost function definition was insensitive to final loop voltage profiles which have a long current diffusion timescale. This meant that these discharges would be relatively close to a stationary state but the final evolution would take considerable time. Therefore the discharge would most likely evolve into a plasma state that would violate some of the plasma constraints. The cost function for stationarity that was added to the possible choices is given by:

$$J_{ss}^{(2)} = ||\sigma_{||} (U_{pl}(\rho) - U_{pl}(\rho = 1))||_{W_{ss}^{(2)}}^2 \quad (6.7)$$

It was found that this definition of the stationarity state cost function is less likely to reach a loop voltage profile with a long diffusion timescale. It is worth noting that a stationary state U_{pl} -profile would be approximately described by:

$$U_{pl}^{ss} \approx \sigma_{||} \frac{\partial\psi}{\partial t} \quad (6.8)$$

Another cost function that was added is motivated by a result from an article by J. Citrin [15]. In this article it was found that the threshold at which a turbulence process known as Ion Temperature Gradient (ITG) mode sets in can be raised by increasing the factor $s(\rho)/q(\rho)$. The threshold is typically described by the behavior given in 6.9 :

$$\frac{|\nabla T_i|}{T_i} \sim 1 + \frac{s(\rho)}{q(\rho)} \quad (6.9)$$

This is schematically illustrated in figure (6.1):

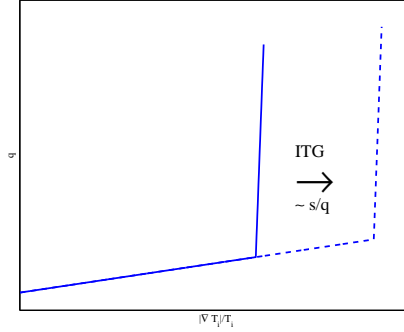


Figure 6.1: Schematic representation of the improvement of the ITG threshold

It is therefore favorable for a hybrid scenario in ITER to have an as broad as possible q -profile at the end of the ramp-up phase. We therefore wish to maximize the volume-averaged factor $\frac{s}{q}$. This is represented by a cost function that has a slightly different form as the other cost functions:

$$J_{s/q} = - \int_0^1 V' \frac{s(t_f, \rho)}{q(t_f, \rho)} d\rho \quad (6.10)$$

6.1.2 Constraints

The operation of ITER is, of-course, limited by constraints on the specifications of the machine. The limitations have to be properly taken into account in order to reach feasible actuator trajectories after the optimization. These constraints have to be translated to constraints on the actuator trajectories and the (predicted) plasma profiles during operation. These constraints are formulated as constraints on the 1D profiles, not on the coil currents or power supply specifically. These constraints typically have 2 different forms:

- Direct linear (in-)equality constraints on the actuator trajectories. These are typically in the form:

$$A_{eq}p = b_{eq} \quad (6.11)$$

$$A_{ineq}p \leq b_{ineq} \quad (6.12)$$

where the actuator trajectories are parameterized by a vector p . One example would be a limitation on the ramp-up speed of the plasma current: $\frac{\partial I_{pl}}{\partial t}$.

- Constraints on the (predicted) plasma profiles. These typically show non-linear behavior with respect to the parametrization of the actuator

trajectories. A constraint is usually described by a function $c_i(t, x(t)) \leq 0$. The full constraint is normally casted into an integral form:

$$\mathcal{C}_i = \int_{t_0}^{t_f} (\max(0, c_i(t, x(t))))^2 dt - \epsilon \leq 0 \quad (6.13)$$

where $\epsilon > 0$ is a small relaxation parameter.

Different motivations for the relevant constraints have been identified during the work for this thesis. Some were already used in the original work, others were updated to resemble ITER relevant values and yet others were introduced during the work on this thesis to accommodate ITER simulations. A list of all the constraints used in the optimization results is given here:

- **Constraint on lowest q:**

The hybrid scenario feature q-profiles which are slightly above one. This way it avoids sawtooth crashes which can trigger NTM's or disruptions. This is imposed as a constraint on q or ι :

$$q(\rho, t) > 1 \quad \forall t, \rho \quad (6.14)$$

- **Constraints on I_p :**

Due to limits on the ITER coils the plasma current is limited but also the ramp-up and ramp-down speed of the plasma current is limited. These constraints, for ITER, are given by:

$$0.5\text{MA} \leq I_p \leq 15\text{MA} \quad (6.15)$$

$$-0.4\text{MA}/s \leq \frac{\partial I_p}{\partial t} \leq 0.25\text{MA}/s \quad (6.16)$$

- **Constraint on $l_i^{(3)}$:**

The ITER team uses an normalized internal inductance denoted by $l_i^{(3)}$. It is given by:

$$l_i^{(3)} = \frac{2V\langle B_p^2 \rangle}{(\mu_0 I_p)^2 R} \quad (6.17)$$

This internal inductance is limited from above by the vertical plasma positioning provided by the coils. Therefore the operational limits on the coils can be translated to a physical upper limit constraint on the internal inductance. The MHD stability of the plasma puts a constraint on the lower limit of $l_i^{(3)}$. The complete constraint for ITER is given by:

$$0.65 \leq l_i^{(3)} \leq 1.2 \quad (6.18)$$

- **Constraint on P_{ECCD} :**

The energy source for the ECCD sources are limited which puts a constraint on the total input power by ECCD sources:

$$\sum_i P_i^{ECCD} \leq 20 \text{ MW} \quad (6.19)$$

- **Constraint on P_{NBI} :**

Similarly, the energy source for the NBI is also limited, putting a separate constraint on the total input power from the NBI sources.

$$\sum_i P_i^{NBI} \leq 33 \text{ MW} \quad (6.20)$$

6.1.3 Local Minima

A problem for non-linear optimization problems is that the obtained minimum of the cost function might not be a global minimum for the parameter space investigated. Mathematically it is very hard to prove that a global minimum even exists but proving that the obtained minimum is the global minimum is even harder to prove. A method to alleviate this problem and reducing the chance of reaching a local minimum at the end of the ramp-up was described in the previously mentioned thesis by F. Felici [10]. This method is based on increasing the parameter space of free parameters and will be briefly reviewed here for the reader.

The procedure followed to minimize the chance of not ending in a global minimum is inspired by multi-grid methods. It starts by choosing a minimal number of optimization parameter, which in our case is typically one free parameter per actuator (apart of course from the fixed NBI actuator). The optimization problem is solved for this case and the solution is taken as the starting point for a new iteration. In this iteration a new time point is added to increase the number of free optimization parameters per actuator. Since the starting point is chosen to be the solution of the previous iteration, the algorithm can always return to the previous solution if no better solution is available. This procedure is repeated until the time grid is divided in enough grid points such that the cost function is no longer decreasing significantly upon increasing the number of free optimization parameters.

An additional problem regarding local minima is caused by the initial conditions. One important point is that the initially chosen actuator trajectory can affect the optimized trajectory. A way to avoid this is by computing a normal distributed set of initial trajectories from which the optimization is done for a relatively small number of iterations. From this set a new set of the best optimized trajectories that differ significantly are then taken to be optimized further for a larger number of iterations.

6.2 L-mode Optimization Results

Now that the optimization set-up has been discussed we are ready to start applying this optimization procedure to the case of ITER simulations. The objective of this research was to see whether this method of optimizing the actuator trajectories yields better results than the more experience based heuristic optimization done in literature. In order to do a comparison between these different optimization methods we need to take a reference case from literature. For our comparison in the L-mode optimization we choose the recently published results by D. Hogewij *et al.* [16]. This reference case and the results obtained in this article are discussed in the next section. After this the numerical optimization results are discussed and compared to this reference case in the following sections. Subsequently the optimal trajectories sensitivity to changes in the model, as well as the active constraints during the ramp-up, are investigated.

6.2.1 Reference Case

The previously introduced reference case reported two different results on the optimization of ITER hybrid scenario. The first optimization results were obtained by using NBI, ECCD and LHCD auxiliary heat and current drive sources. The second results were obtained with only NBI and ECCD sources. Both these results only involve the L-mode confinement regime during the entire simulation. Here we have chosen to only look at the second set of results because the LHCD is not contained in the ITER baseline design and not (yet) simulated in RAPTOR. The article also reported on differences in the actuator trajectories between different transport models. For our purposes it is enough to look only at the previously discussed Bohm-gyroBohm transport model. These results were taken as a starting point and reference case for our numerical optimization. This reference case is characterized by a plasma current ramp-up until $t = 80$ sec, an off-axis NBI source of 16 MW and two ECCD sources located at $\rho \sim .4$ & $.55$ with a combined power of 20

MW. A time trace of the actuator trajectories can be found in figure (6.2):

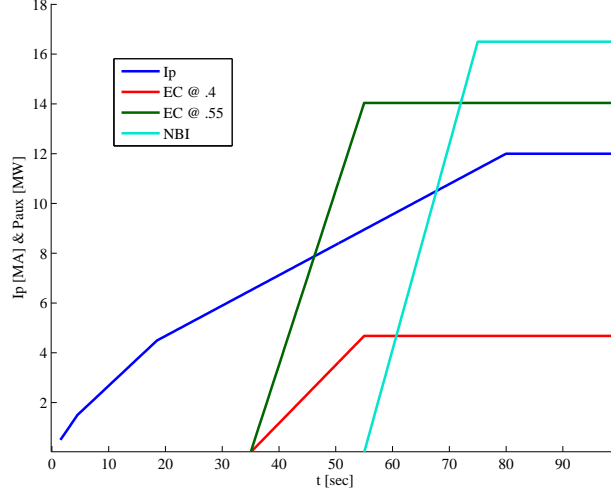


Figure 6.2: A schematic picture of the actuator trajectories in the reference case. Blue line is the plasma current, ECCD sources at $\rho \sim .4$ (red line) and $\rho \sim .55$ (green line) and the NBI source (cyan line)

The specifications of the simulation, both hardware and plasma specifications, were taken from the ITER modeling team predictions on the ITER L-mode hybrid scenario. These specifications include but are not limited to the plasma current ramp rate, n_e and Z_{eff} evolution and the geometry. These specification were taken from the simulation results as an input to the RAPTOR simulation. The evolution of the geometry is a exception in the sense that it is kept fixed in RAPTOR as previously explained. We chose therefore to start the RAPTOR simulation at $t = 20$ seconds because the last closed magnetic surface reaches it final form at this time. The goal of the optimization in [16] was to reach an as broad as possible q-profile at the end of the ramp-up phase. The authors did not however pay attention to the degree of stationarity at the end of the ramp-up. In our optimization we therefore did the optimization for two different cost functions. First was with a cost function that best resembles the goal of the reference case: broadening the q-profile. To this purpose we used the cost function $J_{s/q}$ introduced in previous section (eq. 6.10) as our optimization goal. The second optimization was done with an extra cost function contribution, namely: the degree of stationarity reached at the end of the ramp-up. This was done by using a composite cost function given by:

$$J = \nu_{ss} J_{ss} + \nu_{s/q} J_{s/q} \quad (6.21)$$

The reason to do this is because the fact that stationarity is not reached at the end of the ramp-up could mean that the evolution can cause problems during the flattop phase. These problems could be that the constraints are violated in the flat-top (such as $q < 1$) or that the improved results on the broadening of the q -profile vanish by evolution. The results from the numerical optimization in RAPTOR are shown in the next section.

6.2.2 Optimized Trajectories

Here we will discuss the results from the numerical optimization procedure as discussed in the previous sections and compare their results to the reference case. While obtaining these results we used the procedure of increasing the number of free parameters but the results shown here are the final results, the results at intermediate stages will be discussed in a different section. The time grid points chosen are:

$$t = (20\ 30\ 40\ 50\ 60\ 70\ 80\ 90\ 100) \quad (6.22)$$

As mentioned before we did two different optimizations with two different cost functions. We will first look at the results that had the same objective as the reference case. Subsequently we will look at the results where the objective of increasing stationarity was also introduced. Finally the difference in the resulting trajectories will be shown and discussed.

Optimized volume averaged s/q : broad q -profile

The first case will try to match the objective of the author in the reference case as closely as possible and achieve better results. There are two ways to track our progress with respect to our objectives. The first is comparing the resulting q -profiles visually. This is the way the authors assessed their progress. The second is by comparing the value of the cost function in both RAPTOR and CRONOS. This is a more direct and decisive way of comparing but also slightly more complicated. The complication in comparing the results in this way lies in the fact that the optimal value for the cost function relies on the magnetic equilibrium. For most of the cost functions described in the first section the optimal value is zero. For instance the cost function for stationarity is zero when the plasma reaches a stationary state. This is however not so easy with the cost function $J_{s/q}$. For RAPTOR, because the equilibrium is fixed, the optimal value for $J_{s/q}$ is also fixed. The optimal value and the corresponding q -profile can be calculated by solving an optimization procedure for the available base functions for the different plasma quantities. The optimal value for $J_{s/q}$ in RAPTOR was determined at $J_{s/q} = -253.15$

and the optimal q -profile can be seen in figure (6.3). The same is not true however for CRONOS and this calculation is unfortunately not possible in the CRONOS simulation code. Additionally the optimal value will also differ for each run because the magnetic equilibrium will change which makes a comparison between a RAPTOR and a CRONOS simulation unreliable when using the cost function values. However the comparison of cost function values is still a reliable tool to measure the increase in the ITG threshold, even in CRONOS.

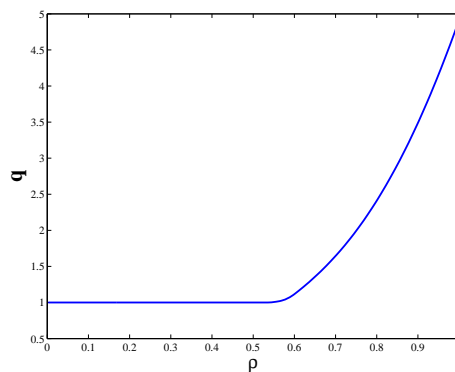


Figure 6.3: *The optimal q -profile with largest $\langle s/q \rangle$ for equilibrium used in RAPTOR as a function of ρ . Whether this profile is physically feasible is not known*

During the work on the optimization it was found that an increase in $J_{s/q}$ could be obtained using the same sources as in the reference case. However it became clear that even better results could be obtained when a more central ECCD source was added to the optimization. Here we will only look at the results from this optimization because they were significantly better than with just two sources. First we will look at the differences in the actuator trajectories. Figure (6.4a) shows the difference in the plasma current ramp-up and figure (6.4b) shows the different ECCD trajectories:

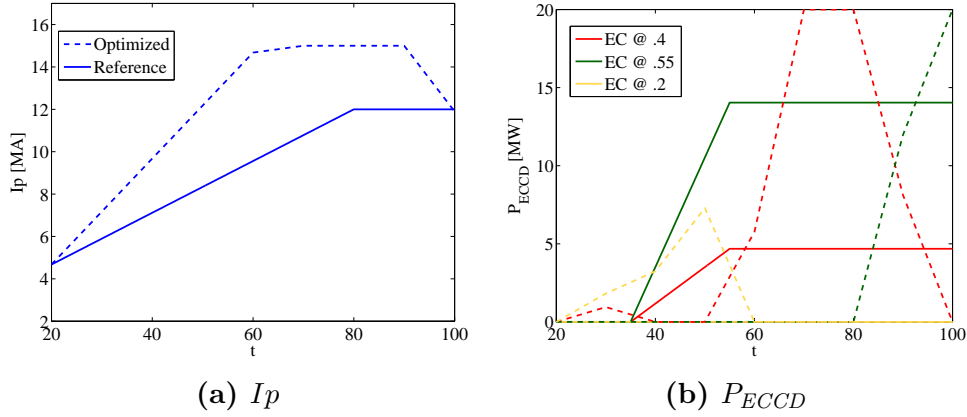


Figure 6.4: The differences in actuator trajectories from the reference case (solid lines) and the numerical optimization (dashed lines)

It can be seen in these picture that for the I_p a current overshoot is favorable for obtaining a broad q -profile at the end of the ramp-up phase. This is in accordance with results found experimentally at JET (Hobirk *et al.* [17]) and with previous results from numerical optimization for TCV. When looking at the ECCD sources is can be seen that the power deposition is moved from the center of the plasma towards more off-axis near the end of the ramp-up phase.

It is also very interesting to look at the improvement in the cost function and to look at the q -profile at the end of the ramp-up. In table (6.2a) the cost function values are shown for the different RAPTOR and CRONOS runs. For RAPTOR there is an extra quantity called $relJ_{s/q}$ defined as:

$$relJ_{s/q} = \frac{J_{opt} - J_{s/q}}{J_{opt} - J_{s/q,ref}}. \quad (6.23)$$

Which gives a measure of the improvement on the distance towards the optimal value of the cost function. It can be seen from this table that cost function is decreased significantly in both RAPTOR and CRONOS. However we cannot speak about the decrease in distance from the optimal value in CRONOS since the optimal value is not known.

Additionally it is very interesting to look at the resulting q -profile as simulated by CRONOS. To this end we have done a CRONOS simulation with the resulting actuator trajectories from the numerical optimization. Figure (6.5) shows a comparison between the q -profile at the end of the ramp-up between the reference case and the numerically optimized case:

	RAPTOR $J_{s/q}$	rel $J_{s/q}$	CRONOS $J_{s/q}$
Reference case	-146.14	1	-215.52
Optimized case	-230.93	0.208	-289.10

Table 6.1: The cost function values for the numerical optimization and the reference case, as calculated by CRONOS and RAPTOR

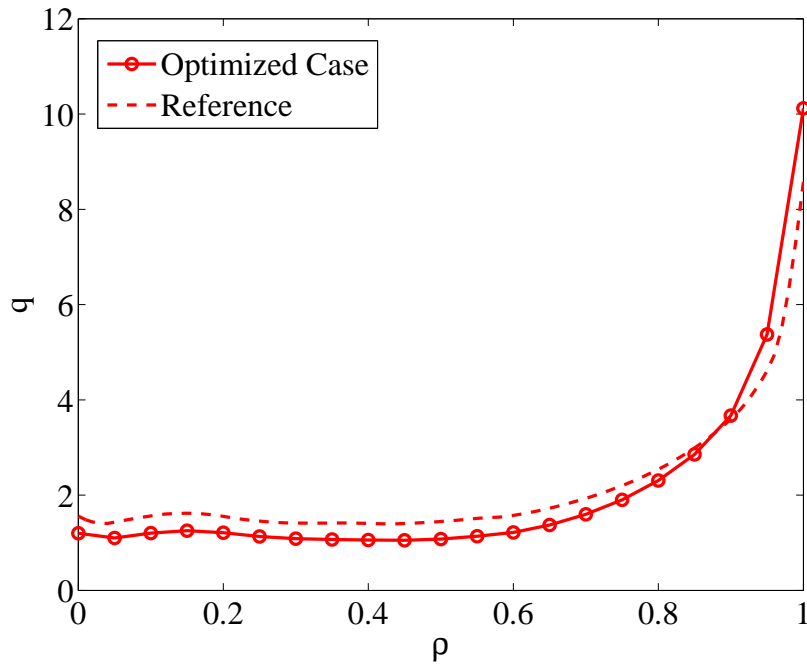


Figure 6.5: Comparison of resulting q -profile as simulated by CRONOS by reference case and with optimized actuators

It is also interesting to look at other profiles at the end of the ramp-up and compare these to the results from RAPTOR. This can be seen in figure (6.6). Here it can be seen that again these seem to match quite well.

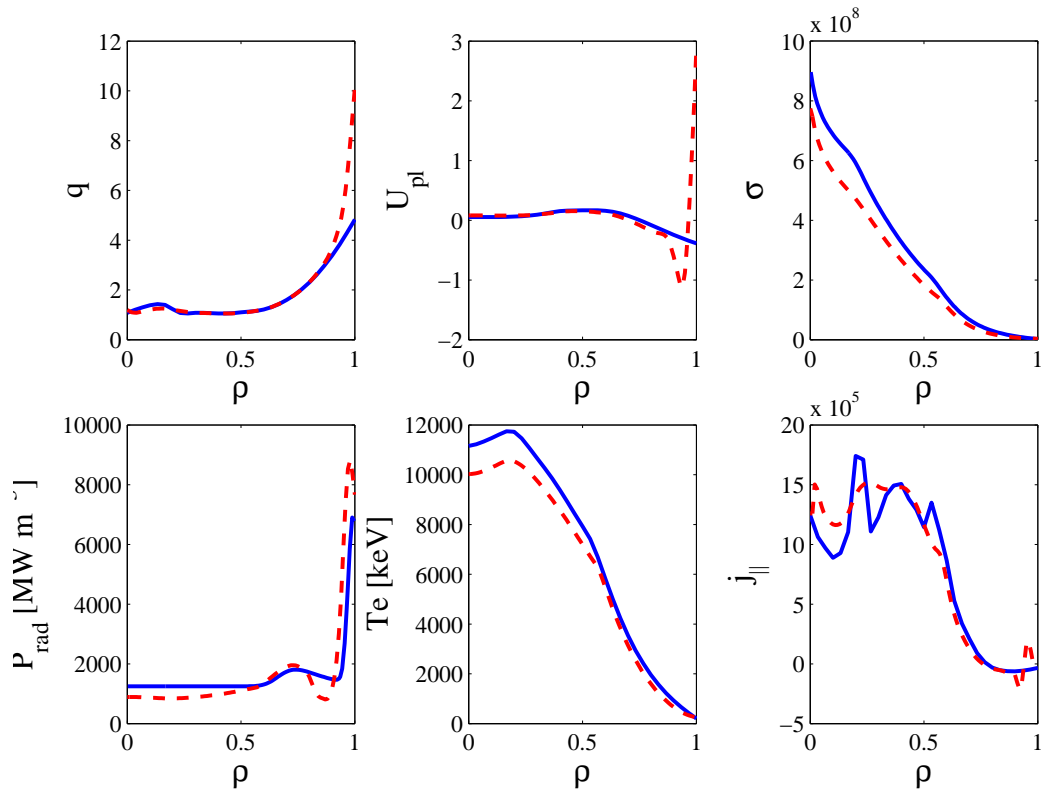


Figure 6.6: Comparing resulting profiles from CRONOS (red dashed line) and RAPTOR (blue solid line) at the end of the ramp-up phase

Adding Stationarity

Now we have seen that it is possible to significantly decrease the cost function using numerical optimization, when looking at a cost function that aims to broaden the q -profile, it is time to also seek to improve on the degree of stationarity. This is done, as mentioned previously, by adding a term to the cost function looking to flatten the U_{pl} -profile. The same numerical optimization is then done with this new cost function. Figures (6.7a) and (6.7b) show the resulting actuator trajectories compared to the reference case.

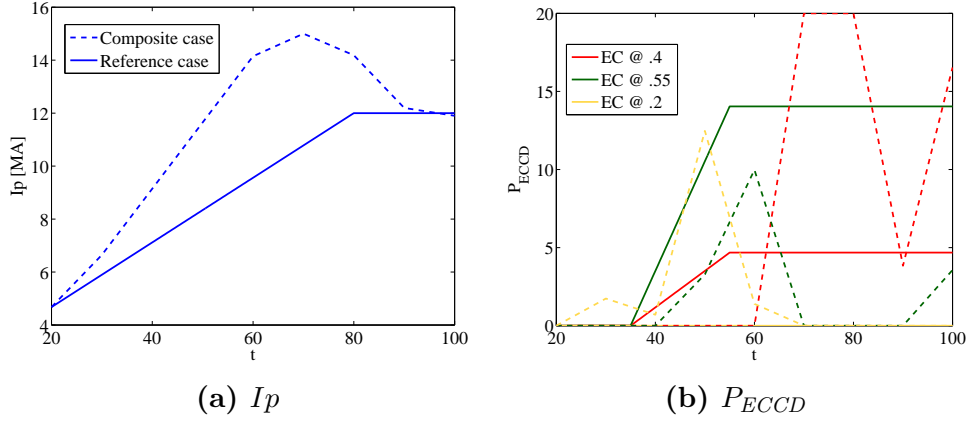


Figure 6.7: The differences in actuator trajectories from the reference case (solid lines) and the numerical optimization (circles)

A table with the results for the different cost functions. The different contributions were weighted using ν_{ss} and $\nu_{s/q}$ such that their contributions have the same order of magnitude. The table shows the unweighed contributions in order to compare with previous results.

	RAPTOR $J_{s/q}$	rel $J_{s/q}$	CRONOS $J_{s/q}$
Reference case	-146.14	1	-215.52
$J_{s/q}$ case	-230.93	0.208	-289.10
Composite case	-218.29	0.326	-257.61

(a) The $J_{s/q}$ values.

	RAPTOR J_{ss}	rel J_{ss}	CRONOS J_{ss}	rel J_{ss}
Reference case	8.993×10^{-3}	1	7.958×10^{-3}	1
$J_{s/q}$ case	0.7546	8.392	2.6437	332.20
Composite case	3.126×10^{-4}	0.035	9.67×10^{-4}	0.122

(b) The J_{ss} values.

Table 6.2: The unweighed cost function values for the optimized case and the reference case, as calculated by CRONOS and RAPTOR. This optimization was done for a composition cost function: $J = \nu_{ss}J_{ss} + \nu_{s/q}J_{s/q}$

Again, additional insight into the improvement of the optimization can be gained by looking at the new q -profile as simulated by CRONOS. The q -profile at the end of the ramp-up phase is shown in figure (6.8a) where it is compared to the reference case and the previously optimized case were only

$J_{s/q}$ was optimized. Since now we have also tried to optimize the U_{pl} -profile it is also interesting to look at this for the different trajectories. Figure (6.8b) shows the U_{pl} -profile for the reference case, the $J_{s/q}$ optimized case and the case were $J = \nu_{ss}J_{ss} + \nu_{s/q}J_{s/q}$ was optimized.

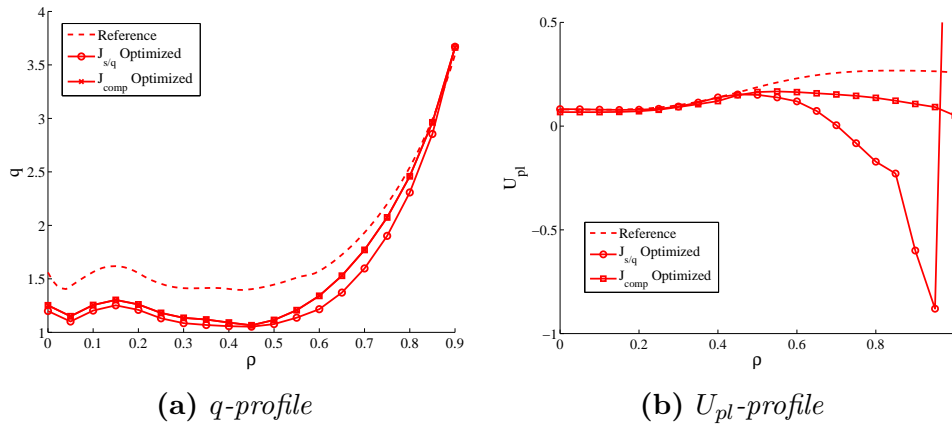


Figure 6.8: A comparison of the q and U_{pl} -profiles at the end of the ramp-up for the reference case (dashed lines), the $J_{s/q}$ optimized case (circles) and the composite case (squares).

Just as in the previous case it is also interesting to look at the differences in the resulting profiles from RAPTOR and CRONOS. Figure (6.9) shows the most relevant profiles as calculated by the two different simulation codes:

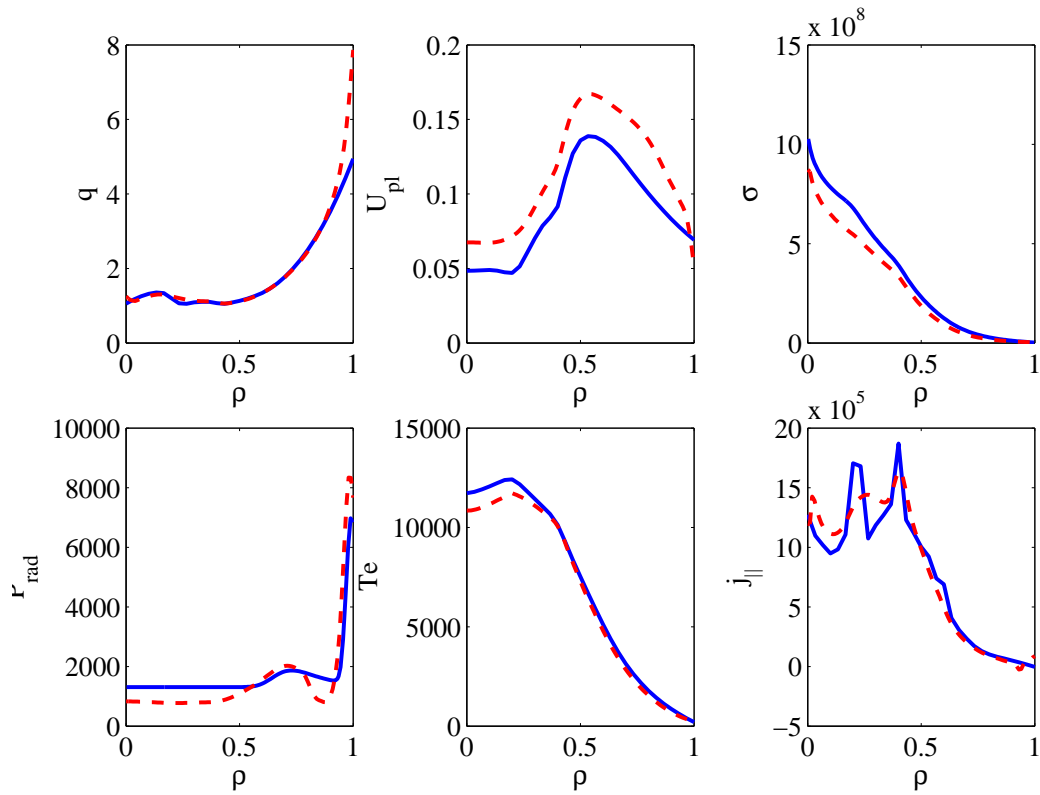


Figure 6.9: Comparing RAPTOR and CRONOS resulting profiles at the end of the ramp-up phase when using the optimized actuators

It is clear from figure (6.4) and (6.7) that the resulting actuator trajectories are significantly different when optimizing for different cost functions. It is interesting to see where these trajectories differ from each other. Figure (6.10) shows a comparison of the two different trajectories. It can be seen that the biggest change is in the trajectory of the most off-axis ECCD source. This is also the general picture in other L-mode optimization that were done during the work on this thesis. The more off-axis the ECCD source, the more sensitive it is to changes in the composition of the cost function.

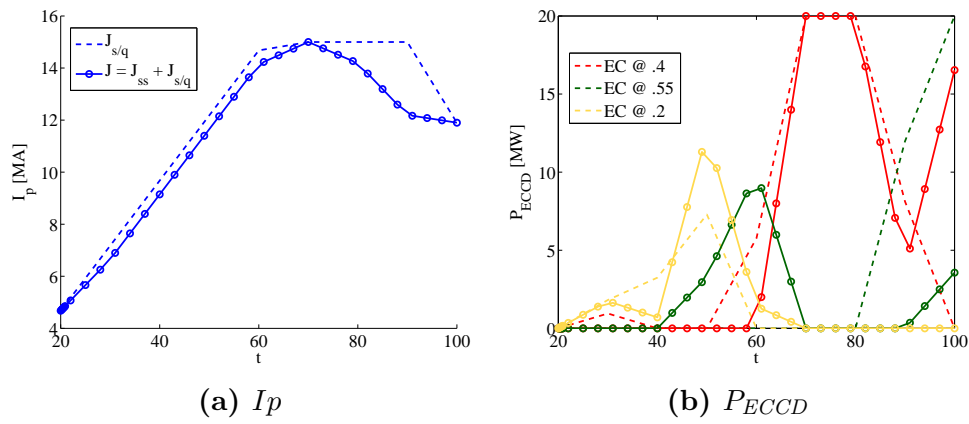


Figure 6.10: The differences in actuator trajectories from the $J_{s/q}$ optimization (dashed lines) and the composite optimization (circles)

A different way to show the differences between the optimized cases and the reference cases is given here. Figure 6.11 shows the actuators trajectories, a contour plot of the q and T_e -profiles. Lastly the final time q , U_{pl} and T_e -profiles are shown to show the improvement of the different objectives.

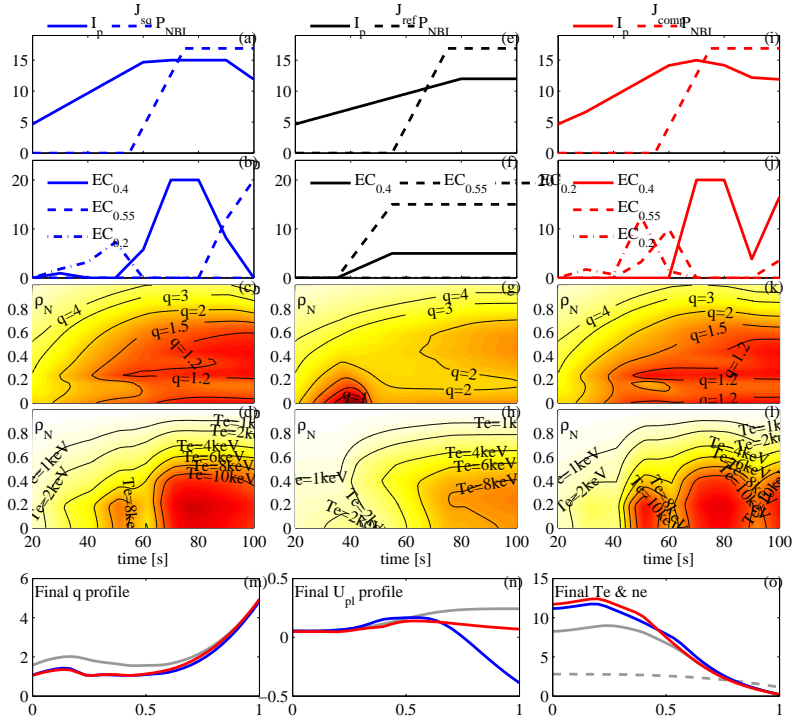


Figure 6.11: Comparison of the results from RAPTOR simulations of the optimized actuator trajectories, contour plot of the q and T_e profiles and the q , T_e and U_{pl} profiles at $t = t_f$. First column (gray lines) shows the reference case, second column (blue lines) show results from J_{sq} optimization and the last column (red lines) are from J_{comp} optimization.

6.2.3 Model Sensitivity

The optimization done in the above pictures were all based on both theory and empirical based models for the conductivity, transport and many other aspects of tokamak physics. These models are however not yet tested for the ITER tokamak since it is still under construction. The resulting trajectories for this specific model might not be generalizable for a wider range of model parameters. This problem is encountered in a wider range of computational sciences and it is important to investigate to what extent the solutions are appropriate for other model parameters. To this extent we do additional simulations of the optimal trajectories found in the optimization but now with different model parameters. These new model parameters are perturbations of the original model parameters with a gaussian distribution of 5 – 25% for the different parameters. The resulting profiles from these new simulations are analyzed and the renewed cost function values are calculated. The distri-

bution of the new cost function values gives information about how general the optimal trajectory is with respect to the model. In this analysis we have chosen to perturb the following model parameters:

- $\sigma_{\parallel}^{\text{pert}} \sim 90\% - 110\% * \sigma_{\parallel}^{\text{old}}$
- $a_{gB}^{\text{pert}} \sim 75\% - 125\% * a_{gB}^{\text{old}}$
- $a_B^{\text{pert}} \sim 95\% - 105\% * a_B^{\text{old}}$
- $Z_{eff}^{\text{pert}} \sim 95\% - 105\% * Z_{eff}^{\text{old}}$

As mentioned before, measures are taken to minimize the chance of reaching a local minima by increasing the number of free time grid points per actuator. Here this method can be used to identify the optimal number of free points while still reaching a result that is generalizable to a wider range of model parameters. This can be seen in figure (6.12) where, for the optimal trajectory at each number of free points a set of 30 new model parameters are recalculated. This figure shows that the nominal cost function value decreases with increasing number of free time grid points. It also shows that with increasing number of free points the spread of cost function values decreases until 6 free times points is reached, after which the spread increases again. The combined effect of a decrease in nominal cost function and spread after perturbations at the 6th free time point make this the optimal number of free points. This is however a choice which can be different when other purposes are investigated. Because the decline in nominal cost function after 5 free points might also be related to the specific time point added at that step in the optimization procedure, it is probable that this decline is not a generic feature of the optimization. An optimization with a different adding sequence, so adding the different time grid points at different iterations of the optimization procedure, lead to the discovery that this is indeed not a generic feature of the optimization.

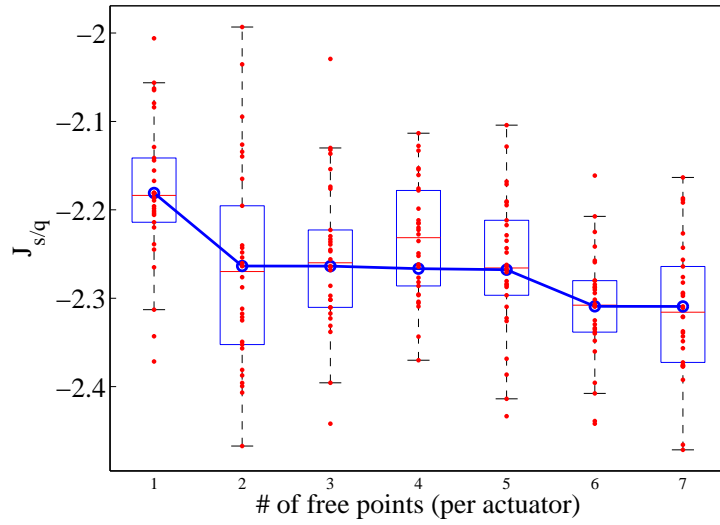


Figure 6.12: Cost function values for increasingly larger number of parametrization points. The solid blue line shows the nominal $J_{s/q}$ decreasing with increasing number of free time points. Also shown are the cost function values for simulations with perturbed model parameters (boxplot and red dots). The boxes show the range of 25% – 75% from the mean value of the distribution.

6.2.4 Active Constraints

In this section we look at which constraints limit the further improvement of the cost function.

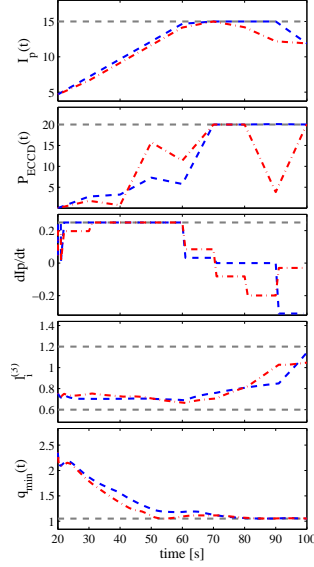


Figure 6.13: Blue dashed lines show the time traces of the constraints as simulated for the J_{sq} optimized actuators. The dashed dotted red lines show the constraint as simulated for the J_{comp} optimized actuators. The dashed green lines show the implemented constraints

6.3 H-mode Optimization

The optimization in the L-mode shows an improved result over the heuristic method of optimization because the complexity of the actuator trajectories considered is far greater with the numerical optimization. Therefore this method considers trajectories normally not considered with heuristic optimization. To give a complete overview of the capabilities of this method of optimization we also like to show optimized trajectories for a H-mode hybrid scenario. To this extend we show here the results from optimization for the 3 different cost function compositions when employed in the H-mode discharge simulation. As a starting point for the H-mode optimization we simulated a baseline scenario which resembles the typical ITER hybrid scenario envisioned by the ITER modeling team. To this extend the transport model was adapted, as mentioned previously, to simulate the H-mode pedestal. The pedestal height is set by using a control method at $T_{e,ped} \sim 5 \text{ keV}$ and the pedestal width is fixed at $\rho_{ped} = 0.9$. Additionally the electron density n_e parametrization was adapted to resemble the L-H transition and simulate the typical H-mode evolution. Using this density parametrization and the

modifications in the transport model we set up a baseline H-mode simulation. The goal of this baseline scenario is to reach H-mode hybrid scenario with $T_{e0} \sim 20 \text{ keV}$, $n_{e0} \sim 9.5 \cdot 10^{19}$ and $Q \sim 5$ while using full NBI (33 MW) and ECCD power (20 MW) and satisfying $q > 1$. The results from this baseline scenario are shown in figure 6.14:

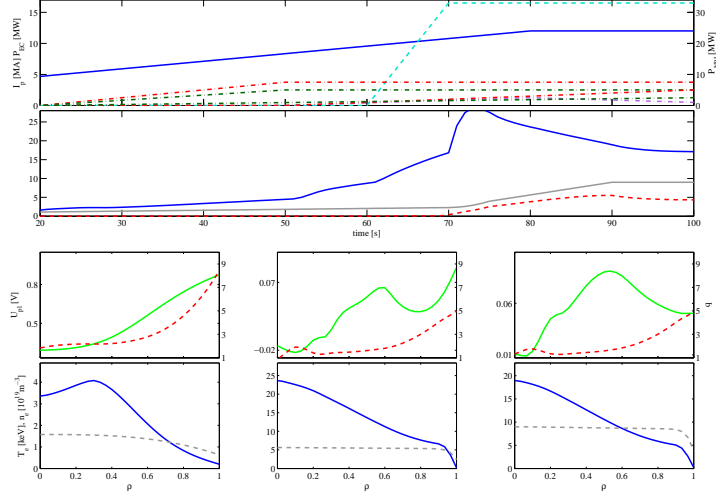


Figure 6.14: The first row shows the time traces of the actuator trajectories and second row the core electron temperature and density. Figure (c-h) show the U_{pl} , q , T_e and n_e -profiles at 3 different times

This baseline scenario was used as a starting point for the H-mode optimization. The same procedure was used as for the L-mode optimization and 3 different cost functions were considered: $J_{s/q}$, J_{ss} and $J_{comp} = \nu_{ss} J_{ss} + \nu_{sq} J_{sq}$. The weights are chosen such that the optimization of J_{comp} reaches an intermediate state between J_{ss} and J_{sq} . The results of the optimized trajectories are shown in figure 6.15. This figure shows the I_p actuators, contour plots of $V P_{ECCD}(\rho, t)$, q and T_e and final time q , T_e and U_{pl} -profiles.

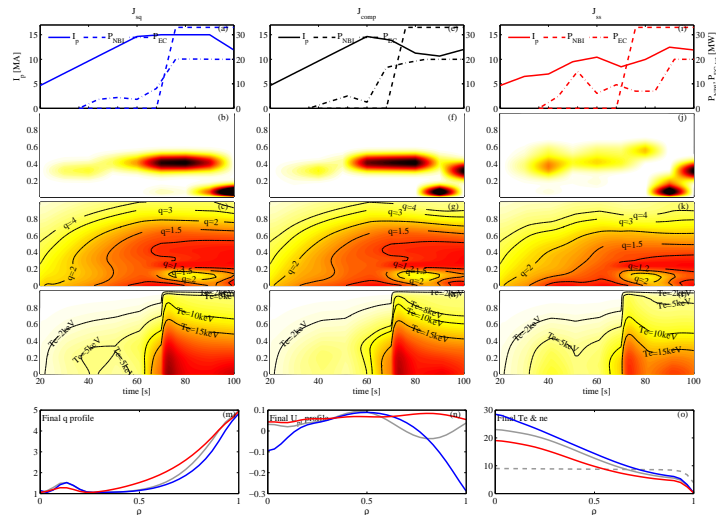


Figure 6.15: The first row show the I_p and NBI actuators. The second row figures show the contour plots of V^*P_{ECCD} , the third row figures show the q contour plots and the fourth row show the T_e contour plots. The last row shows the final time plots of the q , U_{pl} and T_e profiles of J_{sq} (blue lines), J_{ss} (red lines) and J_{comp} (green lines)

Again the time traces of the constraints give extra insight in the limitations on the optimization. The time traces plots are shown in figure (6.16)

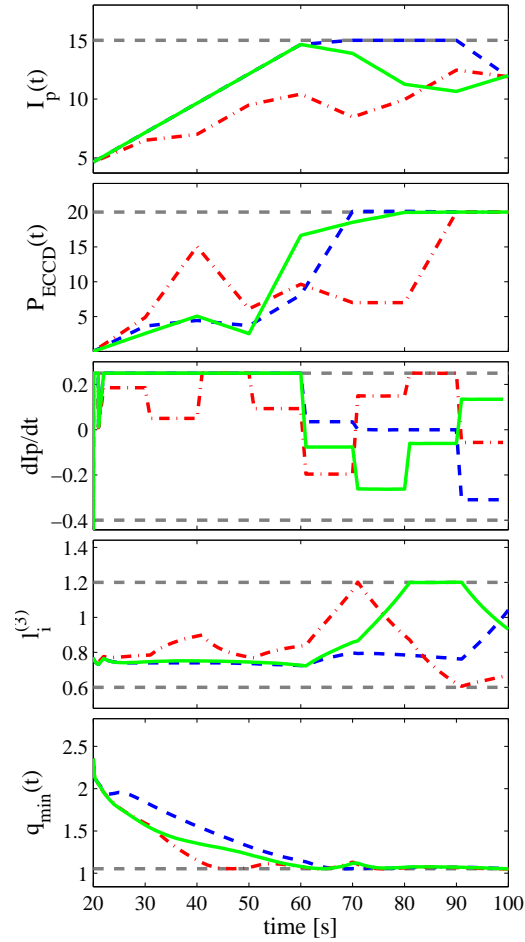


Figure 6.16: *time traces of the implemented constraints for the different optimized actuators simulated. The J_{sq} constraints are in blue dashed lines, the J_{ss} in red dashed dotted lines and J_{comp} in solid black lines. The dashed green lines show the implemented constraints*

In H-mode the model sensitivities were also investigated by evaluating the cost function values for 30 model perturbations. Figure 6.17 shows the distribution of the cost function values after simulating the optimized trajectories at each intermediate step with the perturbed model.

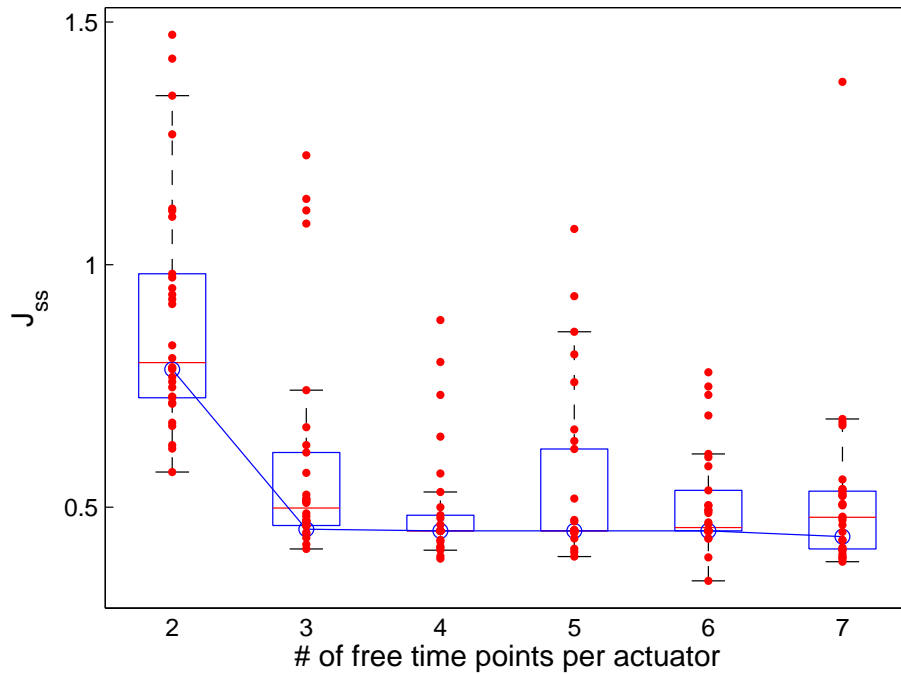


Figure 6.17: shows the nominal (blue line) decrease with increasing number of free time points. Additionally it shows the cost function values when simulating the optimized trajectory with a set of 30 different perturbed model parameters (red dots). The boxes show the 25th and 75th percentile interval

In addition to the model perturbations we also did the full optimization procedure with different n_e evolution and L-H transition timing. We altered the global density by a multiplication factor of both $n_e = 0.9 \cdot n_{e,old}$ and $n_e = 1.1 \cdot n_{e,old}$. The L-H transition was shifted by 10 seconds in both direction: $t_{LH} = 60$ and $t_{LH} = 80$. Figure (6.18-6.19) show the differences between the original optimization and the optimization for these perturbed scenarios.

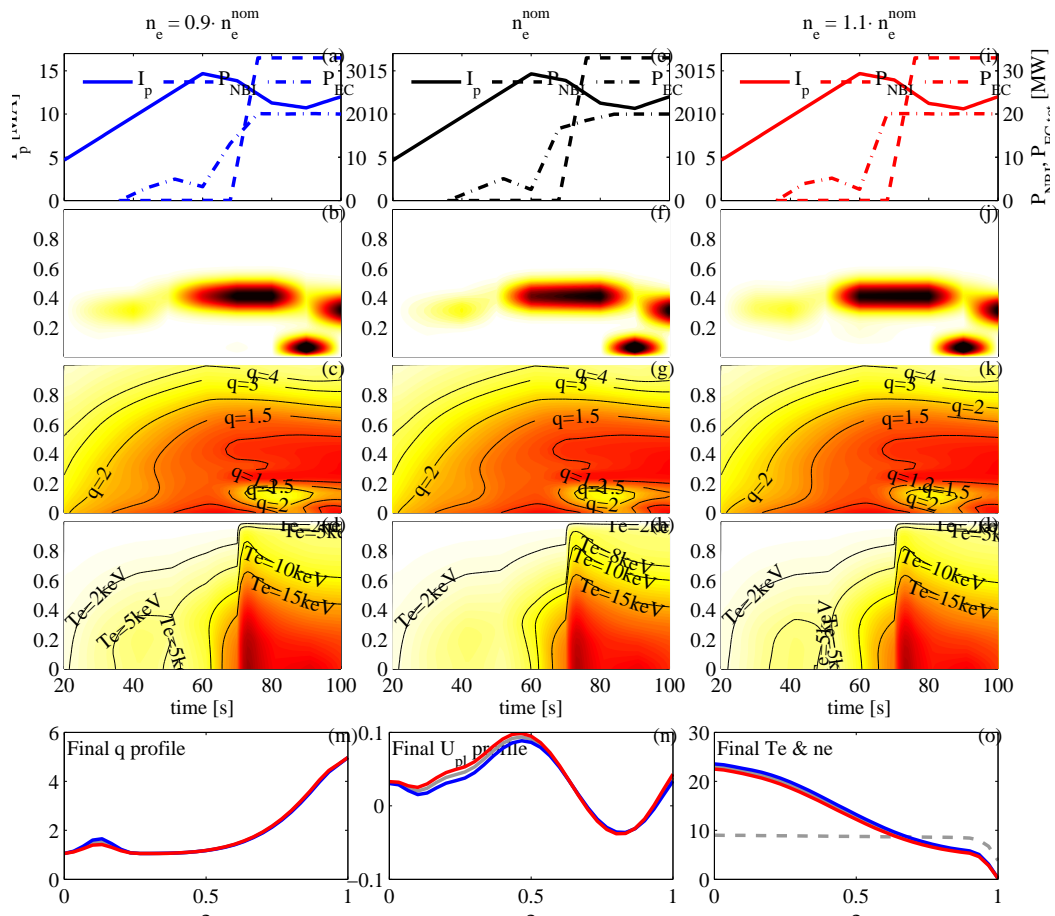


Figure 6.18: shows the differences between the nominal optimization (gray lines), an optimization with $n_e = 0.9 \cdot n_e^{nom}$ (blue lines) and with $n_e = 1.1 \cdot n_e^{nom}$ (red lines)

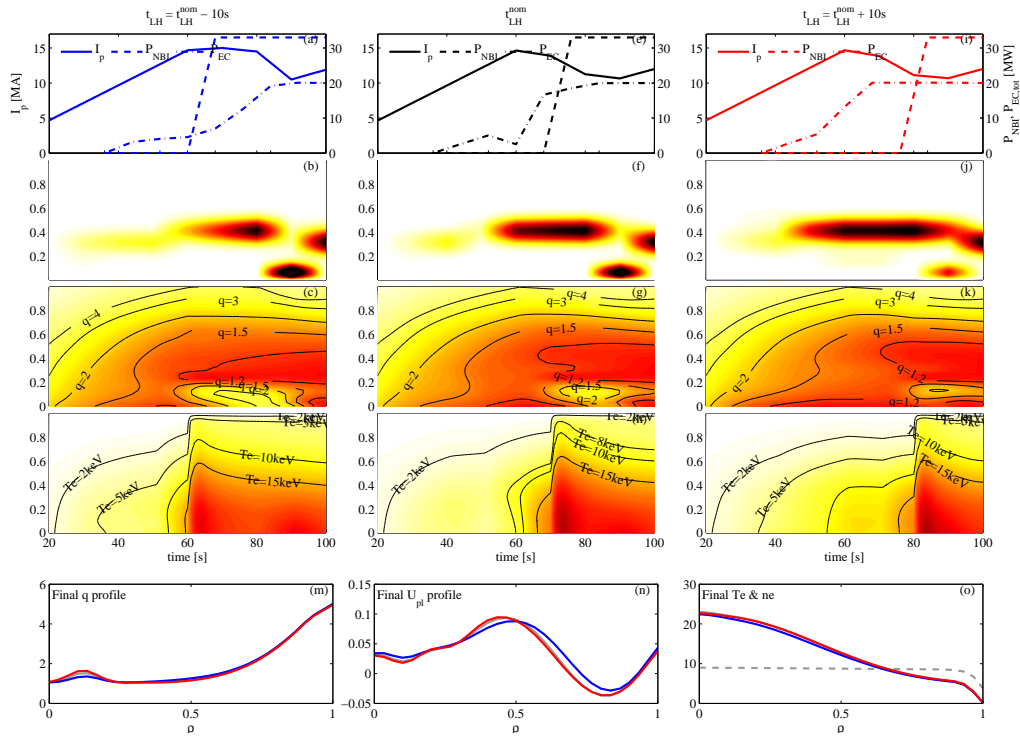


Figure 6.19: shows the differences between the nominal optimization $t_{LH} = 70\text{sec}$ (gray lines), an optimization with $t_{LH} = 80\text{sec}$ (blue lines) and with $t_{LH} = 60\text{sec}$ (red lines)

Chapter 7

Conclusion

This work has shown that the optimization of actuator trajectories can be successfully done by means of a numerical optimization procedure such as is implemented in the RAPTOR simulation code. The fact that the non-linear plasma physics is in some cases not the full theoretical description does not outweigh the fact the trajectories considered are far more complex than any trajectory ever considered under heuristic optimization. The performance can be a factor of 10 times better as the heuristic optimization as verified by the more complete simulation code CRONOS. It is however still uncertain whether a global or local minima was reached at the end of the optimization. It is however nearly impossible to prove that a global minima is reached in such a complex non-linear optimization problem. Certainly in the case where a composition cost function is employed we know that either of the contributions can not attain its most favorable final time profile.

7.1 Outlook

A great opportunity in improving this work lies in doing the work for the JET and ASDEX-Upgrade tokamaks. The numerical optimization done here for ITER has not been verified experimentally since ITER is still under construction. Applying this optimization tool to existing tokamaks and therefore verifying its performance with experimental data can greatly increase the validity of numerical optimization in the fusion plasma physics community. Further more new procedures can be developed to minimize the chance of reaching a local minimum at the end of the optimization. A few examples would be to start from an ensemble of un-optimized actuator trajectories in order to remove any bias from these begin conditions. An other would be to perturb the optimized trajectories in order to escape a local minimum.

A final point of improvement would be to also do the L-mode optimization with more generic ECCD power sources and to also include counter-ECCD sources or ECRH sources.

7.2 Acknowledgements

Here I would like to put a word of thanks and appreciation towards my daily supervisors: Dick Hogewij and Federico Felici. They have greatly helped me develop my knowledge of plasma physics, plasma control, computer modeling and optimization. They have also shown me what it entails to be a researcher and have given me the chance to discover my path by trial and error while still steering me in the right direction. A word of thanks to Jonathan Citrin is also in place for his help in learning my way with CRONOS and his words of encouragements.

Appendix A

Mathematical Derivations

A.1 Flux Diffusion

$$\begin{aligned}\langle \nabla \cdot \mathbf{F} \rangle &= \frac{\partial}{\partial V} \int (\nabla \cdot \mathbf{F}) dV = \frac{\partial}{\partial V} \int (\nabla \cdot \mathbf{F}) R d\phi dl_p \frac{d\psi}{|\nabla\phi|} \\ &= \frac{\partial}{\partial V} \oint \mathbf{F} \cdot \frac{\nabla V}{|\nabla V|} R d\phi dl_p \quad \text{by Gauss} \\ &= \frac{\partial}{\partial V} 2\pi \oint \mathbf{F} \cdot \nabla V \frac{\partial\psi}{\partial V} \frac{R dl_p}{|\nabla V|} \\ &= \frac{\partial}{\partial V} \langle \mathbf{F} \cdot \nabla V \rangle\end{aligned}\tag{A.1}$$

We can derive for a general scalar field $F(t,x,y,z)$ that defines a scalar function $H(t) = \int_V F dV$, where the integral is taken over a volume enclosed by a flux surface, the following:

$$\left. \frac{\partial H}{\partial t} \right|_{\psi=\text{cst}} = \int_V \frac{\partial F}{\partial t} dV + \oint F \mathbf{u}_\psi \cdot d\mathbf{S}_\psi = \int_V \frac{\partial F}{\partial t} dV + \oint F \mathbf{u}_\psi \cdot \frac{\nabla\psi}{|\nabla\psi|} dS_\psi.\tag{A.2}$$

In order to reach the identities in (2.36) we use the following:

$$\begin{aligned}
\int_V \frac{\partial \mathbf{B}}{\partial t} \cdot \nabla \phi dV &= - \int_V ((\nabla \times \mathbf{E}) \cdot \nabla \phi) dV \\
&= - \int_V \nabla \cdot (\mathbf{E} \times \nabla \phi) dV \\
&= - \oint_S (\mathbf{E} \times \nabla \phi) \cdot \frac{\nabla \psi}{|\nabla \psi|} dS \\
&= - \oint_S \mathbf{E} \cdot (\nabla \phi \times \nabla \psi) \frac{dS}{|\nabla \psi|} \\
&= -2\pi \oint_S \mathbf{E} \cdot \mathbf{B}_p \frac{dS}{|\nabla \psi|}
\end{aligned} \tag{A.3}$$

and by using

$$\mathbf{B} \cdot \nabla \psi = \text{cst} \tag{A.4}$$

we can derive for the toroidal electric field:

$$\begin{aligned}
-\mathbf{B} \cdot \nabla \frac{\partial \psi}{\partial t} &= \nabla \psi \cdot \frac{\partial \mathbf{B}}{\partial t} = -\nabla \psi \cdot (\nabla \times \mathbf{E}) \\
&= \nabla \cdot (\nabla \psi \times \mathbf{E}) \\
&= \nabla \cdot (\nabla \psi \times \mathbf{E}_\phi) \\
&= \nabla \cdot (\nabla \psi \times \nabla \phi R E_\phi) \\
&= -\nabla \cdot (2\pi \mathbf{B}_p R E_\phi) \\
&= -\mathbf{B} \cdot \nabla (2\pi R E_\phi)
\end{aligned} \tag{A.5}$$

Therefore, using the continuity equation for ψ

$$\mathbf{u}_\psi \cdot \nabla \psi = -2\pi R E_\phi \tag{A.6}$$

Last derivation is that of the parallel current term in (2.31):

$$\begin{aligned}
\frac{\langle \mathbf{j} \cdot \mathbf{B} \rangle}{B_0} &= \left\langle \frac{T}{2\pi\mu_0 R^2} \Delta^* \psi - \frac{1}{2\pi\mu_0 R^2} \frac{\partial T}{\partial \psi} (\nabla \psi)^2 \right\rangle \\
&= \frac{1}{2\pi\mu_0} \left(T \left\langle \nabla \cdot \left(\frac{\nabla \psi}{R^2} \right) \right\rangle - \left\langle \frac{\nabla T \nabla \psi}{R^2} \right\rangle \right) \\
&= \frac{1}{2\pi\mu_0} \left(T \frac{\partial}{\partial V} \left\langle \frac{\nabla \psi}{R^2} \cdot \nabla V \right\rangle - \left\langle \frac{\nabla T \nabla \psi}{R^2} \right\rangle \right)
\end{aligned} \tag{A.7}$$

Then using the following relations:

$$\begin{aligned}
\nabla\psi &= \frac{\partial\psi}{\partial\mathbf{x}} = \frac{\partial\psi}{\partial\rho}\nabla\rho, \\
\nabla T &= \frac{\partial T}{\partial\rho}\nabla\rho, \\
\nabla V &= \frac{\partial V}{\partial\rho}\nabla\rho = V'\nabla\rho \quad \text{and} \\
\frac{\partial}{\partial V} &= \frac{1}{V'}\frac{\partial}{\partial\rho}.
\end{aligned} \tag{A.8}$$

We can rewrite the last term in (A.7) into:

$$\cdots = \frac{1}{2\pi\mu_0} \left(\frac{T}{V'} \frac{\partial}{\partial\rho} \left(\left\langle \frac{(\nabla\rho)^2}{R^2} V' \frac{\partial\psi}{\partial\rho} \right\rangle \right) - \left\langle \frac{(\nabla\rho)^2}{R^2} \right\rangle \frac{\partial T}{\partial\rho} \frac{\partial\psi}{\partial\rho} \right) \tag{A.9}$$

$$= \frac{1}{2\pi\mu_0} \left(\frac{T4\pi^2}{V'} \frac{\partial}{\partial\rho} \left(G_2 \frac{\partial\psi}{\partial\rho} \right) - \frac{4\pi^2}{V'} G_2 \frac{\partial T}{\partial\rho} \frac{\partial\psi}{\partial\rho} \right) \tag{A.10}$$

$$= \frac{2\pi R_0}{V'\mu_0} \left(J \frac{\partial}{\partial\rho} \left(G_2 \frac{\partial\psi}{\partial\rho} \right) - G_2 \frac{\partial J}{\partial\rho} \frac{\partial\psi}{\partial\rho} \right) \tag{A.11}$$

$$= \frac{2\pi R_0 J^2}{V'\mu_0} \frac{\partial}{\partial\rho} \left(\frac{G_2}{J} \frac{\partial\psi}{\partial\rho} \right). \tag{A.12}$$

A.2 Particle Diffusion

As said in the main part of this thesis the particle transport equation starts at the continuity equation for a given particle species α (2.46). Next we take the volume integrated form of the continuity equation over a volume enclosed by a toroidal flux surface:

$$\int \frac{\partial n_\alpha}{\partial t} dV + \oint n_\alpha \mathbf{u}_\alpha \cdot \nabla\Phi \frac{dS}{|\nabla\Phi|} = \int s_\alpha dV \tag{A.13}$$

Now we can use (A.2) to write this as:

$$\frac{\partial}{\partial t} \Big|_\Phi \int n_\alpha dV + \oint_S n_\alpha \mathbf{u}_\Phi \cdot \nabla\Phi \frac{dS}{|\nabla\Phi|} + \oint n_\alpha \mathbf{u}_\alpha \cdot \nabla\Phi \frac{dS}{|\nabla\Phi|} = \int s_\alpha dV \tag{A.14}$$

Now using the fact that $\nabla\rho \parallel \nabla\Phi$:

$$\frac{\partial}{\partial t} \Big|_\Phi \int n_\alpha dV + \oint_S n_\alpha (\mathbf{u}_\Phi + \mathbf{u}_\alpha) \cdot \nabla\rho \frac{dS}{|\nabla\rho|} = \int s_\alpha dV \tag{A.15}$$

Using the flux averaging definition:

$$\frac{\partial}{\partial t} \Big|_{\Phi} \int \langle n_{\alpha} \rangle V' d\rho + V' \langle n_{\alpha} (\mathbf{u}_{\Phi} + \mathbf{u}_{\alpha} \cdot \nabla \rho) \rangle = \int V' \langle s_{\alpha} \rangle d\rho. \quad (\text{A.16})$$

When we take the radial derivative of this expression we get:

$$\frac{\partial}{\partial t} \Big|_{\Phi} (\langle n_{\alpha} \rangle V') + \frac{\partial}{\partial \rho} \underbrace{(V' \langle n_{\alpha} (\mathbf{u}_{\Phi} + \mathbf{u}_{\alpha}) \cdot \nabla \rho \rangle)}_{\Gamma_{\alpha}} = \langle s_{\alpha} \rangle V' \quad (\text{A.17})$$

Using the definition of the time derivative derived in (2.34) we conclude this derivation:

$$\frac{1}{V'} \left(\frac{\partial}{\partial t} - \frac{\rho B_0}{2B_0} \frac{\partial}{\partial \rho} \right) (\langle n_{\alpha} \rangle V') + \frac{1}{V'} \frac{\partial}{\partial \rho} \Gamma_{\alpha} = S_{\alpha} \quad (\text{A.18})$$

Appendix B

Bohm-GyroBohm Transport Model

This appendix chapter is dedicated to the motivation and mathematical expressions of the Bohm-GyroBohm transport model as reported in the original paper by Erba *et al.* [11]. The basis of their model is the fact that the coefficient of the thermal diffusion in plasmas can be written as:

$$\chi = \chi_0 F(x_1, x_2, \dots) \quad (\text{B.1})$$

where χ_0 has the correct dimensions and the function F is a function of dimensionless plasma parameters. Most standard choice for χ_0 is the usual Bohm expression B.2:

$$\chi_0 = c \frac{T_e}{eB_t}. \quad (\text{B.2})$$

The dimensionless plasma parameters that enter the function F are the plasma normalized gyro-radius ρ^* , the plasma pressure β , safety factor q and the normalized scale length L_p . These are given in expressions (B.3)-(B.5):

$$\rho^* = \frac{m_i^{1/2} T_e^{1/2}}{eaB} \quad (\text{B.3})$$

$$\beta = \frac{8\pi p}{B^2} \quad (\text{B.4})$$

$$L_p = \frac{p}{a^* |\nabla p|} \quad (\text{B.5})$$

When F is proportional to ρ^* the plasma turbulence scale length is of the order of the plasma gyro-radius. In this case the diffusivity depends on local plasma parameters and is said to be gyro-Bohm like. However, when the function F does not depend on ρ^* , the scale length of the plasma turbulence is

in the order of the plasma minor radius. In this case the transport properties are called Bohm like. When these two different types of transport are mixed with appropriate coefficients, they form a transport model that is effective at modeling the transport for a wide range of plasma parameters. When the Bohm term (B.6) and the gyro-Bohm (B.7) are added appropriately they form the Bohm-GyroBohm transport model used in RAPTOR and CRONOS as given in Eq (B.8):

$$\chi_{e,i}^B = \frac{T_e}{B_t L_{pe}} q^2 \quad (\text{B.6})$$

$$\chi_{e,i}^{GB} = \frac{a^* \nabla T_e}{B_t} \rho_i^* \quad (\text{B.7})$$

$$\chi_{e,i} = \alpha_{e,i}^B \chi_{e,i}^B + \alpha_{e,i}^{GB} \chi_{e,i}^{GB} \quad (\text{B.8})$$

Bibliography

- [1] J.A. Wesson. Tokamaks. *Plasma Physics and Controlled Fusion*, 46(3):573, 2004.
- [2] F. L. Hinton and R. D. Hazeltine. Theory of plasma transport in toroidal confinement systems. *Rev. Mod. Phys.*, 48:239–308, Apr 1976.
- [3] D. V. Sivukhin. Coulomb Collisions in a Fully Ionized Plasma. *Reviews of Plasma Physics*, 4:93, 1966.
- [4] H.-S. Bosch and G.M. Hale. Improved formulas for fusion cross-sections and thermal reactivities. *Nuclear Fusion*, 32(4):611, 1992.
- [5] Charles F.F. Karney. Fokker-planck and quasilinear codes. *Computer Physics Reports*, 4(34):183 – 244, 1986.
- [6] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer series in operations research and financial engineering. Springer, New York, NY, 2. ed. edition, 2006.
- [7] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.
- [8] A. Fasoli, C. Gormenzano, H.L. Berk, B. Breizman, S. Briguglio, D.S. Darrow, N. Gorelenkov, W.W. Heidbrink, A. Jaun, S.V. Kononov, R. Nazikian, J.-M. Noterdaeme, S. Sharapov, K. Shinohara, D. Testa, K. Tobita, Y. Todo, G. Vlad, and F. Zonca. Chapter 5: Physics of energetic ions. *Nuclear Fusion*, 47(6):S264, 2007.
- [9] M. Schneider, L.-G. Eriksson, F. Imbeaux, and J-F. Artaud. Self-consistent simulations of the interaction between fusion-born alpha particles and lower hybrid waves in iter. *Nuclear Fusion*, 49(12):125005, 2009.
- [10] F. Felici. *Real-Time Control of Tokamak Plasmas: from Control of Physics to Physics-Based Control*. PhD thesis, EPFL, 2011.

- [11] M. Erba, T. Aniel, V. Basiuk, A. Becoulet, and X. Litaudon. Validation of a new mixed bohm/gyro-bohm model for electron and ion heat transport against the iter, tore supra and start database discharges. *Nuclear Fusion*, 38(7):1013, 1998.
- [12] W. A. Houlberg, K. C. Shaing, S. P. Hirshman, and M. C. Zarnstorff. Bootstrap current and neoclassical transport in tokamaks of arbitrary collisionality and aspect ratio. *Physics of Plasmas*, 4(9):3230–3242, 1997.
- [13] O. Sauter, C. Angioni, and Y. R. Lin-Liu. Neoclassical conductivity and bootstrap current formulas for general axisymmetric equilibria and arbitrary collisionality regime. *Physics of Plasmas*, 6(7):2834–2839, 1999.
- [14] J.F. Artaud, V. Basiuk, F. Imbeaux, M. Schneider, J. Garcia, G. Giruzzi, P. Huynh, T. Aniel, F. Albajar, J.M. An, A. Bcoulet, C. Bourdelle, A. Casati, L. Colas, J. Decker, R. Dumont, L.G. Eriksson, X. Garbet, R. Guirlet, P. Hertout, G.T. Hoang, W. Houlberg, G. Huysmans, E. Joffrin, S.H. Kim, F. Kchl, J. Lister, X. Litaudon, P. Maget, R. Masset, B. Pgouri, Y. Peysson, P. Thomas, E. Tsitrone, and F. Turco. The cronos suite of codes for integrated tokamak modelling. *Nuclear Fusion*, 50(4):043001, 2010.
- [15] J. Citrin, J.F. Artaud, J. Garcia, G.M.D. Hogeweyj, and F. Imbeaux. Impact of heating and current drive mix on the iter hybrid scenario. *Nuclear Fusion*, 50(11):115007, 2010.
- [16] G.M.D. Hogeweyj, J.-F. Artaud, T.A. Casper, J. Citrin, F. Imbeaux, F. Kchl, X. Litaudon, I. Voitsekhovitch, and the ITM-TF ITER Scenario Modelling Group. Optimizing the current ramp-up phase for the hybrid iter scenario. *Nuclear Fusion*, 53(1):013008, 2013.
- [17] J Hobirk, F Imbeaux, F Crisanti, P Buratti, C D Challis, E Joffrin, B Alper, Y Andrew, P Beaumont, M Beurskens, A Boboc, A Botrugno, M Brix, G Calabro', I Coffey, S Conroy, O Ford, D Frigione, J Garcia, C Giroud, N C Hawkes, D Howell, I Jenkins, D Keeling, M Kempenaars, H Leggate, P Lotte, E de la Luna, G P Maddison, P Mantica, C Mazzotta, D C McDonald, A Meigs, I Nunes, E Rachlew, F Rimini, M Schneider, A C C Sips, J K Stober, W Studholme, T Tala, M Tsalas, I Voitsekhovitch, P C de Vries, and JET EFDA contributors. Improved confinement in jet hybrid discharges. *Plasma Physics and Controlled Fusion*, 54(9):095001, 2012.