



Osborne, Christopher M. J. (2021) *Casting a new light on radiative transfer in solar flare models: synthesis and inversion*. PhD thesis.

<https://theses.gla.ac.uk/82584/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>  
[research-enlighten@glasgow.ac.uk](mailto:research-enlighten@glasgow.ac.uk)

# **Casting a New Light on Radiative Transfer in Solar Flare Models: Synthesis and Inversion**

**Christopher M. J. Osborne**  
**BSc (Hons)**

Submitted in fulfilment of the requirements for the  
degree of Doctor of Philosophy

Astronomy and Astrophysics Group  
School of Physics and Astronomy  
University of Glasgow



**University  
of Glasgow**

September 2021

---

This thesis is my own composition except where indicated in the text.  
No part of this thesis has been submitted elsewhere for any other degree  
or qualification.

**Copyright © 2021 Christopher M J Osborne**

1st December 2021

---

*I send you this letter on a falling star. Reentry will score and test it but will not melt it away. I write in fire across the sky, a plummet to match your rise.*

---

**This Is How You Lose the Time War**

— Amal El-Mohtar & Max Gladstone

# Abstract

The chromosphere is a thin layer of the observable solar atmosphere maintained in a complex energy balance that is not well understood. Chromospheric conditions can only be probed through observations of radiation, primarily atomic spectral lines. The shapes and intensities of these spectral lines cannot be correctly reproduced by assuming local thermodynamic equilibrium (LTE) conditions, and are instead affected by non-local radiation fields. During solar flares, an enormous amount of energy is deposited in the chromosphere, heating the plasma, provoking dramatic dynamic reactions and substantial changes in the observed line profiles. A two-pronged approach is therefore required to develop our understanding of this region and its reaction to flares: detailed observations, and a theoretical framework derived from complex modelling by which to interpret them.

In this thesis we focus on techniques for modelling optical spectral lines in solar flare conditions, including a reassessment of previous assumptions, the radiative influence of a flare on neighbouring atmosphere, and present a novel machine learning inversion technique. Our choice of optical lines is motivated by the exceptional resolution and cadence of modern and upcoming ground-based observatories, enabling a highly detailed exploration of this region. All of the radiative modelling presented is performed with our new Python framework *Lightweaver*, which aims to facilitate the development of complex radiative transfer simulations.

In Chaps. 1, 2, and 3, we introduce the necessary background material. We first present a brief description of the outer layers of the solar atmosphere and general introduction to solar flares. Then, in Chap. 2, we provide an overview of field-aligned radiation hydrodynamic modelling of solar flares, and associated numerical techniques, including an in-depth discussion of radiative transfer in non-LTE conditions. In Chap. 3, we introduce the two primary spectral lines used in our research,  $H\alpha$  and Ca II 854.2 nm, along with an overview of the inverse problem of radiative transfer (determining the atmospheric conditions responsible for an observed spectrum), and the concept of response functions,

## Abstract

which describe the radiative response to a perturbation in atmospheric parameters. We also provide a brief introduction to machine learning techniques.

The design philosophy and implementation of *Lightweaver* is described in Chap. 4. We also provide validation examples for the various features of this framework, presenting comparisons with state-of-the-art radiative transfer codes RH and SNAPI, along with two further tests to validate its time-dependent treatment of atomic level populations.

In Chap. 5, we apply *Lightweaver* to the synthesis of radiation from flare models produced by the radiation hydrodynamic code RADYN and present three different case studies. Firstly, using two different RADYN simulations, we investigate the photoionising effects of the hydrogen Lyman lines on Ca II and show that this has a significant effect on the emergent line profiles and a 10–15 % change in chromospheric radiative losses. Secondly, the importance of a time-dependent treatment of Ca II is tested, with only minor deviations found due to ignoring these effects. Finally, we present a modification of this tool to include the effects of partial frequency redistribution (PRD) and synthesise the radiation from a RADYN model, despite the poor convergence of our technique. The Doppler-like approximate PRD treatment used for the Lyman lines is found to remain relatively accurate in the flaring simulation, but the Ca II K line profile and radiative losses differ more significantly. We also present a time-dependent formulation of response functions that, with further development, should allow for greater interpretability of spectral line formation in complex radiation hydrodynamic models.

The extension and application of *Lightweaver* to a two-dimensional slab of quiet Sun plasma adjacent to a RADYN flare model is presented in Chap. 6. This slab is held at a constant pre-flare temperature and density structure, whilst the atomic level populations and electron density are allowed to vary with the incoming radiation. Significant enhancements in the H $\alpha$  and Ca II 854.2 nm line profiles are found in excess of 1 Mm from the flare, although no continuum effects are seen in this simple model. These enhancements found in this simple model are compared against observations from the CRisp Imaging SpectroPolarimeter (CRISP) on the Swedish Solar Telescope (SST) and effects of a similar order of magnitude are found, although substantial differences remain.

RADYNVERSION, described in Chap. 7, is a novel deep learning based inversion technique built on an invertible neural network, allowing inference of chromospheric conditions multiple orders of magnitude faster than conventional techniques, outside of the constraints of statistical and hydrostatic equilibrium that are ill-suited to flares. We discuss in detail the theory of its construction, and provide a brief proof of concept application to SST/CRISP data, which is found to agree with previous analysis of this event.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>List of Figures</b>	<b>ix</b>
<b>Preface &amp; Acknowledgements</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Layers of the Sun . . . . .	2
1.2 Solar Flares . . . . .	5
1.3 Conventions . . . . .	8
<b>2 Numerical Flare Modelling</b>	<b>10</b>
2.1 Radiation Hydrodynamics . . . . .	11
2.2 An Eye to the Future: Radiative Magnetohydrodynamics . . . . .	13
2.3 Solar Flare Heating . . . . .	14
2.4 Introduction to Radiative Transfer . . . . .	15
2.4.1 The Formal Solution . . . . .	16
2.4.2 Short-Characteristics Methods . . . . .	18
2.4.3 Other Formal Solvers . . . . .	19
2.4.4 LTE vs NLTE . . . . .	20
2.4.5 Collisional Rates . . . . .	21
2.4.6 Emissivity and Opacity . . . . .	22
2.4.7 Radiative Rates . . . . .	25
2.4.8 General Source Function . . . . .	25
2.4.9 Iterative Solutions . . . . .	25
2.4.10 Solving the Multilevel NLTE Problem . . . . .	27
2.4.11 Time-Dependent Population Updates . . . . .	30
2.4.12 Partial Frequency Redistribution . . . . .	31
2.4.13 Charge Conservation . . . . .	33
2.4.14 Determining Numerical Convergence . . . . .	35

## Contents

2.5	Introduction to Hydrodynamics and Conservation Laws . . . . .	36
2.5.1	Numerical Approaches . . . . .	38
2.5.2	Riemann Problems . . . . .	40
2.5.3	Godunov’s Method and Higher Order Reconstructions . . . . .	41
2.5.4	Numerical Fluxes . . . . .	42
2.5.5	Time Integration, Stability, and Splitting Schemes . . . . .	43
2.6	Conduction . . . . .	45
2.7	Discussions . . . . .	46
<b>3</b>	<b>Optical Flare Observations and Inversions</b>	<b>48</b>
3.1	Important Optical Spectral Lines . . . . .	48
3.1.1	The Swedish Solar Telescope . . . . .	51
3.1.2	The Daniel K Inouye Solar Telescope . . . . .	52
3.2	Introduction to Inverse Problems . . . . .	52
3.2.1	Milne-Eddington Inversions . . . . .	55
3.2.2	Generalisation Through Response Functions . . . . .	56
3.2.3	Forward Modelling . . . . .	62
3.2.4	In the Context of the Latent Space . . . . .	62
3.3	Introduction to Machine Learning . . . . .	63
3.3.1	Artificial Neural Networks . . . . .	64
3.3.2	General Function Approximations . . . . .	65
3.3.3	Training via backpropagation . . . . .	66
3.3.4	Difficulties training DNNs . . . . .	67
<b>4</b>	<b>The <i>Lightweaver</i> Radiative Transfer Framework</b>	<b>69</b>
4.1	Philosophy . . . . .	69
4.2	Accessibility & Code Overview . . . . .	71
4.3	Model Atoms . . . . .	72
4.4	Other Implementation Details . . . . .	73
4.5	Parallelisation . . . . .	76
4.6	Validation . . . . .	77
4.7	<i>Lightspinner</i> . . . . .	83
4.8	Discussions . . . . .	84
<b>5</b>	<b>Radiative Transfer with</b>	
	<b>Time-Dependent Populations</b>	<b>85</b>
5.1	A Brief Dissection of RADYN and a Possible Future of RHD Modelling . . .	85
5.2	Minority Species Modelling . . . . .	87

## Contents

5.3	Reprocessing RADYN Simulations with the <i>Lightweaver</i> Framework . . . . .	89
5.3.1	Advection . . . . .	91
5.4	Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines . . . . .	93
5.4.1	The RADYN simulations . . . . .	96
5.4.2	Line Profiles . . . . .	97
5.4.3	Radiative Losses . . . . .	101
5.5	Case Study: Is Full Time-Dependence Necessary? . . . . .	104
5.6	Case Study: Partial Frequency Redistribution . . . . .	107
5.6.1	Modifying Our <i>Lightweaver</i> -based Model to Support PRD . . . . .	107
5.6.2	Results . . . . .	108
5.6.3	Discussion . . . . .	110
5.7	Digging Deeper: Time-Dependent Response Functions . . . . .	112
5.8	Discussions . . . . .	119
<b>6</b>	<b>Two-Dimensional Radiative Transfer</b> . . . . .	<b>123</b>
6.1	The Formal Solver in Two-Dimensions . . . . .	123
6.1.1	The BESSER method . . . . .	125
6.1.2	Evaluation Order and Boundary Conditions . . . . .	128
6.1.3	Implementation Details . . . . .	130
6.1.4	Validation . . . . .	131
6.2	2D Simulation Configuration . . . . .	135
6.3	Simulation Results . . . . .	139
6.3.1	Observed Radiation . . . . .	140
6.3.2	Spectroscopic Results . . . . .	143
6.3.3	Population Changes . . . . .	149
6.3.4	Importance of Time-Dependence . . . . .	156
6.3.5	Horizontal Resolution . . . . .	160
6.4	Observational Comparison . . . . .	163
6.5	Discussions . . . . .	169
<b>7</b>	<b>RADYNVERSION</b> . . . . .	<b>171</b>
7.1	The RADYNVERSION Model . . . . .	172
7.2	Training Data . . . . .	176
7.3	Training Method . . . . .	176
7.4	Maximum Mean Discrepancy . . . . .	179
7.5	Validation . . . . .	181
7.6	Proof of Concept Results . . . . .	183

*Contents*

7.7 Discussion . . . . .	187
<b>8 Concluding Remarks</b>	<b>189</b>
<b>Bibliography</b>	<b>197</b>

# List of Figures

1.1	Temperature and total hydrogen number density ( $n_H$ ) structure of the FALC atmosphere model. . . . .	3
3.1	Grotrian (term) diagram of Ca II bound terms of interest. . . . .	49
3.2	Diagram illustrating the degeneracy present in traditional inversions. . . . .	53
3.3	Resolving the degeneracy of inversions through the addition of a latent space. . . . .	54
3.4	Failure of pseudoinverse in inferring quadratic varying source function. . . . .	58
4.1	Comparison of <i>Lightweaver</i> and RH synthesis of Ca II 854.2 nm from the FALC atmosphere with different electron density solutions. . . . .	77
4.2	Comparison of the <i>Lightweaver</i> , RH, and SNAPI synthesis of Ca II 854.2 nm from the FALC atmosphere with complex velocity profile and LTE electron density. . . . .	78
4.3	Comparison of <i>Lightweaver</i> and RH synthesis of Ca II K from a RADYN snapshot comparing the effects of PRD. . . . .	80
4.4	Fig. 4 of Judge (2017). Time-dependent response of hydrogen populations to instantaneous temperature change. . . . .	81
4.5	Validation of <i>Lightweaver</i> 's time-dependent population update scheme. . . . .	81
4.6	Validation of time-dependent charge conservation in <i>Lightweaver</i> . . . . .	83
5.1	Overlap between Lyman transitions and Ca II continua in model atoms used. . . . .	94
5.2	Comparison of Ca II 854.2 nm treatments in the F9 simulation. . . . .	97
5.3	Comparison of Ca II 854.2 nm treatments in the F10 simulation. . . . .	98
5.4	Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F9 simulation at $t = 11$ s. . . . .	99
5.5	Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F9 simulation at $t = 20$ s. . . . .	99
5.6	Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F10 simulation at $t = 11$ s. . . . .	100

*List of Figures*

5.7	Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F10 simulation at $t = 20$ s. . . . .	100
5.8	Time evolution of the calcium losses in the F9 simulation. . . . .	101
5.9	Time evolution of the calcium losses in the F10 simulation. . . . .	102
5.10	Effect of calcium losses on total radiative losses in the two simulations. . .	103
5.11	Comparison of time-dependent and statistical equilibrium treatment of Ca II 854.2 nm. . . . .	106
5.12	Comparison of the CRD and PRD treatments of the Ly $\alpha$ line in the F9 model. . . . .	109
5.13	Comparison of the CRD and PRD treatments of the Ca II K line in the F9 model. . . . .	110
5.14	Comparison of the chromospheric radiative losses under CRD and PRD treatments of the Ca II K line in the F9 model. . . . .	111
5.15	Comparison of contribution, temperature response, and velocity response functions for the F9 simulation at $t = 11$ s. . . . .	114
5.16	Comparison of contribution, temperature response, and velocity response functions for the F9 simulation at $t = 20$ s. . . . .	115
5.17	Comparison of contribution, temperature response, and velocity response functions for the F10 simulation at $t = 11$ s. . . . .	115
5.18	Comparison of contribution, temperature response, and velocity response functions for the F10 simulation at $t = 20$ s. . . . .	116
5.19	Comparison of time-dependent and statistical equilibrium response functions. . . . .	118
6.1	Diagram of short-characteristics formal solver in two-dimensions. . . . .	124
6.2	Diagram of sweep order for two-dimensional short-characteristics formal solver. . . . .	128
6.3	Validation of 2D formal solver in static FALC atmosphere with periodic $x$ boundary conditions. . . . .	131
6.4	Difference in Ca II 854.2 nm line profiles between 1 and 2D formal solvers. . . . .	132
6.5	Validation of 2D formal solver with complex velocity field, and with fixed $x$ boundary conditions. . . . .	133
6.6	Configuration of the two-dimensional simulation showing the flaring boundary condition. . . . .	136
6.7	Configuration of two-dimensional simulation to produce radiation that would be observed by a slit spectrograph looking across the model flare. . . . .	138
6.8	H $\alpha$ spectral line and variation in the 2D F9 simulation. . . . .	140
6.9	Ca II 854.2 nm spectral line and variation in the 2D F9 simulation. . . . .	141

*List of Figures*

6.10 H $\alpha$ spectral line and variation in the 2D F10 simulation. . . . .	142
6.11 Ca II 854.2 nm spectral line and variation in the 2D F10 simulation. . . . .	143
6.12 H $\alpha$ spectroscopy in the 2D F9 simulation. . . . .	144
6.13 Ca II 854.2 nm spectroscopy in the 2D F9 simulation. . . . .	144
6.14 H $\alpha$ spectroscopy in the 2D F10 simulation. . . . .	146
6.15 Ca II 854.2 nm spectroscopy in the 2D F10 simulation. . . . .	146
6.16 Spectroscopy at $t = 10$ s in the F10 model with $\chi$ boundary conditions set to zero. . . . .	148
6.17 Contribution function, $\chi_{\nu}/\tau_{\nu}$ and relative population changes at $t = 10$ s in the F9 simulation. . . . .	150
6.18 Contribution function, $\chi_{\nu}/\tau_{\nu}$ , and relative population changes at $t = 10$ s in the F10 simulation. . . . .	152
6.19 Radiation field anisotropy and electron density in F10 model at $t = 10$ s. . .	154
6.20 Comparison of statistical equilibrium and time-dependent treatments at $t = 10$ s for the H $\alpha$ line in the F9 model. . . . .	156
6.21 Comparison of statistical equilibrium and time-dependent treatments of H $\alpha$ as a function of $\chi$ position. . . . .	157
6.22 Comparison of statistical equilibrium and time-dependent treatments of Ca II 854.2 nm as a function of $\chi$ position. . . . .	158
6.23 Comparison of statistical equilibrium and time-dependent treatments of H $\alpha$ with electron density taken from time-dependent model. . . . .	159
6.24 Comparison of statistical equilibrium and time-dependent treatments of Ca II 854.2 nm with electron density taken from time-dependent model. . .	159
6.25 Ratio of outgoing radiation using 41 and 81 horizontal points at $t = 5$ s in the F9 simulation. . . . .	161
6.26 Ratio of outgoing radiation using 41 and 81 horizontal points at $t = 10$ s in the F9 simulation. . . . .	162
6.27 Ratio of outgoing radiation using 41 and 81 horizontal points at $t = 15$ s in the F9 simulation. . . . .	162
6.28 Ca II line core observation of M1.1 flare at time 2014-09-06T17:09:26 . . . .	163
6.29 Extent of enhancement in Ca II 854.2 nm for selected region. . . . .	164
6.30 Offset between the enhancement of the blue and red wings of the Ca II 854.2 nm line and the line core for the labelled region. . . . .	165
6.31 Offset between the enhancement of the blue and red wings of the H $\alpha$ line and the line core for the labelled region. . . . .	165
6.32 Offset between the enhancement of the blue and red wings of the Ca II 854.2 nm line and the line core in the simulations. . . . .	167

*List of Figures*

6.33	Offset between the enhancement of the blue and red wings of the H $\alpha$ line and the line core in the simulations. . . . .	167
7.1	Structure of the invertible RADYNVERSION network. . . . .	174
7.2	Validation of RADYNVERSION forwards model. . . . .	182
7.3	Validation of the RADYNVERSION reverse model. . . . .	182
7.4	CRISP context image of 2014-09-06 M1.1 flare just after flare peak. . . . .	184
7.5	Line profiles to invert, on- and off-ribbon. . . . .	185
7.6	Inversion results for the on-ribbon pixel. . . . .	186
7.7	Inversion results for off-ribbon pixel. . . . .	186

## Preface & Acknowledgements

It took me a long time to *grok* the concepts of NLTE radiative transfer, probably due to hubris. I never imagined, at the start of this endeavour, that it would come to represent the core of my research interests. At some point after learning and playing with LTE methods, I convinced myself that the problem couldn't get much more difficult. After all, how hard can it be to determine the opacity of a plasma? The answer, it turns out, is “quite hard”, but that's a topic for the rest of this thesis. The methods we will discuss typically take a journey through a complex space to arrive at their destination — there is an unfortunate lack of direct methods in this field. I don't know whether I've arrived at my destination yet, but this is certainly a milestone, and there are many people I am grateful to for guiding me along this meandering path.

First and foremost, I would like to thank my supervisor, Lyndsay Fletcher, for letting me explore and follow the excitement gradient into the most interesting little valley. In equal parts always ready to point me in the right direction, tear my bad ideas to shreds, and reassure me that I hadn't just wasted the last  $n$  months.

There is no chance that I would be where I am today without the support and guidance of Paulo Simões who really took me under his wing, introduced me to many aspects of numerical modelling, and was always willing to spend time sounding out new ideas. You really helped me learn how to be a scientist and engage with the community.

To my RT and inversion mentor, Ivan Milić: Thank you! Your enthusiasm is infectious, and your hospitality fantastic. I'll never forget cycling around Boulder, or the “quick questions” that turned into chats that lasted for hours. Thank you to everyone else at the National Solar Observatory for being so welcoming, full of different perspectives, and fun to be around, both in and out of the office. Hazel and Jørgen, thanks for providing me with a home in Boulder, rather than just a place to stay!

Denizens of the 604 office, I don't know half of you as well as I should like... I know that you will all support each other as you finally get back in there. Would that we had had more time in that little box of science: I have missed the whiteboard sessions, the impromptu

## *Preface & Acknowledgements*

chess, and the general banter. John and Aaron, thanks for listening to my regular rants about numerical methods; I hope some of my Python advice has been useful! To everyone else along the astronomy corridor, thank you for your passion and willingness to share the latest exciting discovery, scientific or otherwise, especially at 4 pm coffee.

A personal thank you to Mats Carlsson for taking time to explain the intricacies of RADYN, and nudge me in the directions of interesting problems. Thanks also to Petr Heinzl and Jana Kašparová for listening to my ideas and then ensuring that I test them with sufficient rigour!

To the bois of FLaD, your entertaining company got me through the lockdowns with some amount of (in)sanity intact! And where would I be without the soundscapes that have infused throughout this work? Thanks to Devin Townsend, Eluveitie, Leprous, Myrkur, Babymetal, Nightwish, Wardruna, and Heilung... to name but a few.

I must of course thank my parents for always encouraging my curiosity, making me consider the world around me, and supporting all my endeavours. You made me a person capable of handling this.

Throughout all the time spent at home, one person has stayed at my side (admittedly sleeping most of the time). Thanks for the company Augustus, some cats have earned co-authorships for less. Finally, Clara, none of this would have happened without you. Thank you, pup, for all of your love, understanding, friendship, and encouragement. Let's see where the next adventure takes us.

# 1

## Introduction

### To be taught, if we are fortunate

The Sun is obviously responsible for the existence of life on Earth – after all, there would be no solar system without it! The Earth lies within a narrow orbital zone in which the solar irradiance is precisely sufficient to maintain the very specific temperatures required for life as we know it, without cooking or freezing us. Our Sun is but one of *many* similar stars with similar solar systems, and perhaps, similar life. Whilst the planetary temperature is a necessary condition, it is far from sufficient, and the requirements of life are far more complex than a simple balancing act of “power in, power out”. Instead, the *form* of the energy received, and how the atmosphere reacts to it, has a crucial role in the viability of a planet. Indeed, there is a complex and delicate set of photochemical reactions that both maintain the planetary temperature *and* protect its inhabitants from the more dangerous components of the incoming radiation.

During a solar (or stellar) flare certain dangerous components of a star’s radiative output can be dramatically enhanced. For example, the effects of a flare observed from AD Leonis, an M Dwarf star, on an orbiting Earth-like planet were investigated by [Segura et al. \(2010\)](#). The DNA-damaging ultraviolet (UV) radiation increased at the top of the atmosphere by approximately an order of magnitude, which could lead to disastrous consequences on the health and genetic stability of biological lifeforms caught in the path of this oncoming radiation. This would be equivalent to the UV index going from “low” to “extreme” and back again over the course of an hour! The modelling of this event showed that the increase in the more damaging UVB and UVC bands reaching the planet’s surface was greatly diminished thanks to the planet’s oxygen-rich atmosphere, yielding a peak UV dose rate for DNA damage only slightly higher than that on Earth.

## 1.1 The Layers of the Sun

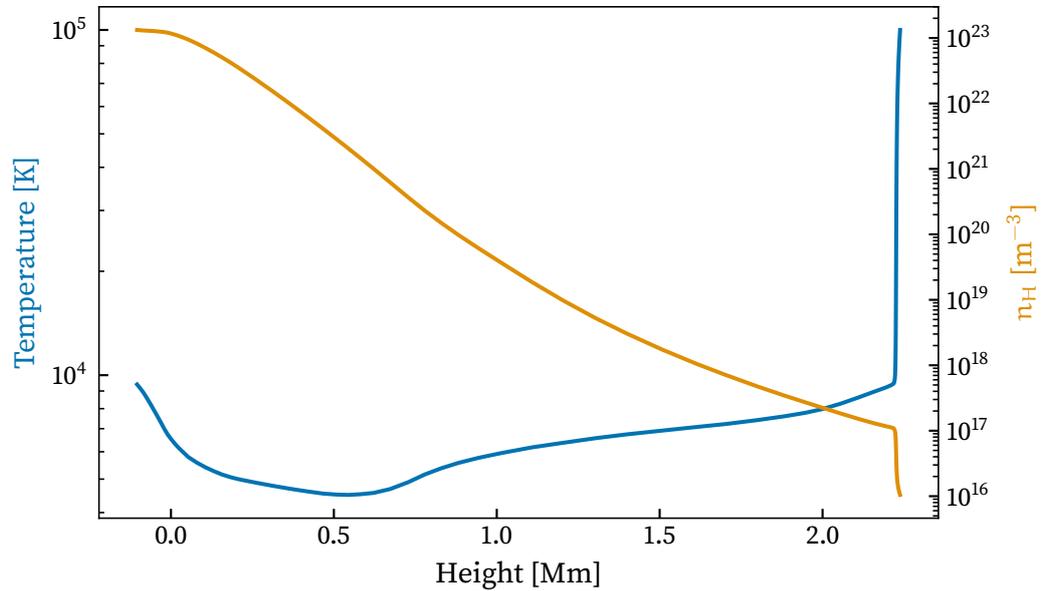
A slightly different atmospheric chemistry could yield different results, but so too could a different flare. Fortunately for us, our solar UV intake varies far less during flares (e.g. [Woods et al. 2006](#)), and the Earth provides an essential envelope that protects, but not with impunity, from the worst this star can throw at us. As our closest star, the Sun most directly affects us, and knowledge gained from this companion can help us understand the environment through which our planet falls, and also the conditions in distant solar systems to which we may one day venture.

Flares produced by our own star have effects clearly visible to life on Earth, from the beautiful aurorae that occasionally adorn the night sky, to disruptive impacts on telecommunications and satellites. Due to the proximity and significance of the Sun, it is the only star, that human science is aware of, to be continuously monitored by a fleet of advanced instrumentation, with a corresponding global community of researchers dedicated to unravelling its mysteries. The intensely dramatic and rapid variations of the solar atmosphere that are observed during energetic flaring events are almost as mesmerising as the aurorae they can produce, and developing an understanding of them is a key component of living long-term as a species with this companion star, whilst its proximity enables investigation on a scale that can only be dreamed of for the distant cosmos. These fascinating questions can and should be answered to slake the natural human thirst for scientific comprehension, but are also key stepping stones to answering some of the questions appearing from the fields of plasma and particle physics, in addition to those of exoplanetary viability.

### 1.1 The Layers of the Sun

The observable solar atmosphere is composed of three main layers of plasma: photosphere, chromosphere, and corona. These are all permeated by a complex, structured, time-varying magnetic field. The photosphere is the innermost observable “surface” of the Sun, and is the origin of the daylight with which we see ([Zirin 1992](#)). The photosphere appears granulated, a pattern of light and dark structure evolving slowly, like the surface of a pot of boiling water, due to the convection cells that form within the photosphere and the solar interior below it. These transport plasma that has cooled down deeper into the atmosphere and renew the hotter surface layer. Larger dark features known as sunspots are often visible on the photosphere: these occur in regions of stronger magnetic field which inhibits the local convection and allows the plasma locked in this region to cool. The photosphere serves as the visible surface of the Sun, and is optically opaque,

## 1.1 The Layers of the Sun



**Figure 1.1:** Temperature and total hydrogen number density ( $n_H$ ) structure of the FAL C atmosphere model of (Fontenla et al. 1993).

primarily due to an abundance of negative hydrogen that forms due to the temperature and density of this region. This  $H^-$  ion is formed by the addition of a second electron to a hydrogen atom: it has a low ionisation potential and is a significant component of solar opacity and emissivity across the optical region and beyond (Hubeny & Mihalas 2014). The photosphere emits as an almost ideal black body with a temperature of approximately 5800 K, and continues radially outward for  $\sim 500$  km from the optically opaque layer until we enter the chromosphere (Carroll & Ostlie 2007).

The chromosphere is  $\sim 2000$  km thick, throughout which the density of the solar plasma drops off rapidly, and in standard reference models the temperature of the plasma starts to rise from approximately 4000–4500 K (Vernazza et al. 1981; Solanki 2004). This lower boundary is known as the temperature minimum region. The increase in temperature above this level clearly shows that the atmospheric temperature structure is not controlled by radiation and conduction alone (which would lead to a monotonic decrease), and there must be additional heating mechanisms at work (Gurman 1992). In Fig. 1.1 we show the temperature and total hydrogen density structures of the FAL C semiempirical atmosphere (Fontenla et al. 1993). This model was constructed to attempt to reproduce observed spectral lines and continua which form throughout the solar atmosphere, and a height of 0 Mm corresponds to an optical depth of unity at a wavelength of 500 nm. The

## 1.1 The Layers of the Sun

optical depth defines the number of mean free paths, or average number of scattering events, for a photon between two points, thus an optical depth of one is often considered to be opaque. Due to its low density, very little broadband light can be observed from the chromosphere. Instead, we image this region with narrowband observations of spectral features, primarily absorption and emission lines.

At the top of the chromosphere, and the top of the FAL C model, there is a dramatic increase in temperature (from  $\sim 20\,000$  K to  $\sim 10^6$  K), with a corresponding decrease in density. The average ionisation of the plasma increases significantly and many ultraviolet spectral lines are observed to form in this region. This exceedingly narrow layer, barely a few hundred kilometres thick, is known as the transition region, as it controls the atmosphere's transition from the denser inner layers to the hot and tenuous corona.

The fully-ionised corona extends outwards to a distance of several solar radii, and maintains a high temperature, from which extreme ultraviolet (EUV) and X-ray emission can be observed. The corona presents a complex structure, containing loops of different scales and morphologies, due to closed magnetic field lines, as well as coronal holes where the plasma in a region is ejected outwards into the heliosphere by open magnetic field lines. An obvious, and long-standing, problem in solar physics is the so-called coronal heating problem (reviewed by [Klimchuk 2006](#)): by what mechanism, or *mechanisms*, does the corona reach and maintain its multi-million K temperature? The current consensus ([De Moortel & Browning 2015](#), and references therein) is that this heating must occur through a variety of processes which is likely to include magnetic reconnection (a restructuring of the magnetic field to a lower potential, liberating the stored energy) and the dissipation of magnetohydrodynamic waves.

Despite its much lower temperature than the corona, the chromosphere requires a much larger energy flux to maintain its temperature structure, due to the much higher density of this region ([De Moortel & Browning 2015](#); [Carlsson et al. 2019](#)). The mechanisms proposed for this include acoustic waves propagating from the photosphere for heating the lower chromosphere, and magnetic field effects in the upper chromosphere. The magnetic field can contribute to heating in a multitude of ways including the dissipation and mode conversion of waves, as well as more direct heating through the release of magnetic stresses by reconnection and Joule heating ([Carlsson et al. 2019](#)). In addition to the heating problem, there are numerous other unanswered questions connected to the chromosphere, including but not limited to the mediation of energy during solar flares, the formation of spicules (small dense jets observed in the chromosphere), the transport

## 1.2 Solar Flares

of the magnetic field and waves, the formation of prominences and filaments (cool and dense material rising from the photosphere into the corona).

The solar atmosphere, at least from the upper photosphere, must be treated as a single unit, and thus, problems such as that of the corona's temperature cannot be considered separately. Some of the corona's energy will be conducted downwards into the chromosphere, a process aided by the magnetic field, and by the same, the magnetic field that permeates the corona originates from deeper in the solar atmosphere and has to pass through the chromosphere. Determining the energy balance throughout the atmosphere is a very active field of research, necessitating both complex numerical models and detailed observational results (e.g. [Carlsson et al. 2019](#)).

## 1.2 Solar Flares

Solar flares are intense short-duration flashes of radiation that occur in the solar atmosphere and can be observed across the entire electromagnetic spectrum. The first documented scientific observation of a solar flare was that of [Carrington \(1859\)](#), which occurred spontaneously during a routine observation of a sunspot cluster. This event was so energetic that the ensuing solar storm and magnetic disturbances wrought havoc on the telegraph network, and aurorae were seen as far south as Colombia ([Moreno Cárdenas et al. 2016](#)). The Carrington Event, as it is known, was a particularly spectacular event, and whilst no event of equivalent magnitude has been recorded since, during the peak of the 11 year solar cycle it is not uncommon to observe multiple smaller flares over the course of a day.

Flares occur in active regions, which are areas where a twisted magnetic field in the kG range emerges from the photosphere (for a review of active region evolution see [van Driel-Gesztelyi & Green 2015](#)). The magnetic field in these regions often leads to the generation of sunspots, although the presence of sunspots is not necessary for a flare. Here, we will briefly describe the evolution of a flare and some of the observational characteristics. For a full review of the latter the reader is directed to the works of [Benz \(2008\)](#) and [Fletcher et al. \(2011\)](#). This description follows the *standard model* of flares, often known as the CSHKP model after the authors of the papers involved in its development: [Carmichael \(1964\)](#); [Sturrock \(1966\)](#); [Hirayama \(1974\)](#); [Kopp & Pneuman \(1976\)](#). This model considers a single magnetic field loop or simple arcade of loops, and whilst flares are typically observed in

## 1.2 Solar Flares

regions of more complex geometry with multiple intersecting and neighbouring loops, it correctly reproduces many observed characteristics.

The evolution of a flare can be roughly split into two sections: impulsive phase, and gradual decay. As the potential energy stored in the magnetic field of the active region increases due to twisting and shearing, it can be seen in photospheric magnetograms and can sometimes be visualised by emission from the energetic plasma trapped within it. The photospheric motions and emergence of this magnetic flux builds and stores energy in coronal loops. In the minutes preceding a flare, small-scale brightenings may be observed in the soft X-ray (SXR) and EUV bands. The flare proper, and the impulsive phase, are presumed to begin with a magnetic reconnection event that releases a large amount of magnetic energy previously stored in the coronal magnetic field. In some situations, an amount of material previously held in place by the closed magnetic field lines is carried by an expanding field and ejected along field lines that are now open to the heliosphere. This process is known as coronal mass ejection. At the looptop, near to the reconnection site, small SXR sources are observed due to the heating and acceleration of electrons. Less frequently, non-thermal hard X-ray (HXR, photon energies above 10 keV) sources may also be observed in this region (Krucker et al. 2008). The accelerated electrons then spiral along the magnetic field lines producing radiation in the synchrotron family.

A large portion of the energy released travels down the loop and arrives at the chromosphere, producing extreme heating, rapid ionisation, and driving significant plasma motions. There is a large spike observed in the X-ray flux, notably HXR bremsstrahlung from high-energy non-thermal electrons accelerated by the energy released by the flare colliding with the denser chromosphere. This has traditionally been described by the collisional thick target model (Brown 1971; Hudson 1972), which presumes that these non-thermal electrons are accelerated in the corona. Due to the heating that ensues from this process the chromosphere undergoes evaporation, where the dense heated material starts to rise along the magnetic loop. These motions can be observed in the Doppler shift of spectral lines, and regions of heating can be seen in the enhancement of chromospheric spectral lines, such as the H $\alpha$  line. In a real-world system with neighbouring loops the reconnection process may travel along the loops, triggering these heating processes in turn and leading to ribbons along their footpoints visible in these lines. SXR observations have revealed compact hot sources with short durations (~1 minute) in flare footpoints (Hudson et al. 1994). Analysis of these regions in both SXR and EUV have revealed densities close to those expected at the top of the chromosphere but with temperatures in the millions of K (Mrozek & Tomczak 2004; Graham et al. 2013; Simões et al. 2015).

## 1.2 Solar Flares

Another proposed mechanism for transporting the energy from the coronal reconnection site to the chromosphere is Alfvén heating (Emslie & Sturrock 1982; Fletcher & Hudson 2008), which suggests that energy is transported by Alfvén waves<sup>1</sup> to the denser layers at the base of a flaring loop where electrons are locally accelerated. One argument for this is the so-called “number problem” which expresses the difficulty in resupplying the acceleration site in the low density coronal plasma with sufficient electrons to reflect the number inferred from HXR observations of their chromospheric impact (e.g. Simões & Kontar 2013). The denser regions closer to the electrons’ impact location do not suffer from this same scarcity of electrons, and thus represent a possible resolution to this problem.

After the energy involved in particle acceleration has been expended the flare enters what is known as the gradual decay phase. Due to the lack of particle acceleration the HXR flux drops rapidly, but the plasma continues to evolve hydrodynamically whilst cooling radiatively to slowly relax into a quieter configuration.

Whilst the total radiated energy may seem like a natural metric by which to classify flares, this is an extremely difficult figure to estimate (e.g. Milligan et al. 2014). The magnitude of a flare is instead estimated by the total solar flux between 0.1 and 0.8 nm as measured by the Geostationary Operational Environmental Satellite (GOES) network. These are then classified based on the peak flux, with the smallest class, A, having a peak flux in the range  $10^{-8}$ – $10^{-7}$  W m<sup>-2</sup>. The following classes, in order of increasing flux, are B, C, M, and X, with bins increasing by 1 dex per class, and X having no upper bound. A numerical suffix is appended to these classes representing the measured flux as a multiple of the bin’s lower bound. The frequency of flaring events is found to be inversely proportional to their magnitude, with X-class flares being quite rare (only a handful of events per year), while less energetic events occur far more commonly. Whilst the most energetic events are rare, their distribution is far from uniform. For example, in September 2017, at the end of solar cycle 24, the Sun produced a series of flares from two active regions, including four X-class events, in the space of less than a week.

The primary unanswered questions in flare physics are related to the storage, release, transport, and dissipation of energy, as well as the interaction of the magnetic field of the magnetic field with these. These phenomena modify the atmosphere and lead to directly observable effects, such as changes in spectral signatures, which are our only means for determining the functioning of these mechanisms.

---

<sup>1</sup>These are transverse waves that propagate along the magnetic field (e.g. Tandberg-Hanssen & Emslie 1988).

### 1.3 Conventions

In this thesis we focus on the numerical modelling of solar flares, paying close attention to the formation and interpretation of chromospheric spectral lines. These spectral lines are accessible to ground-based instrumentation and have already yielded great insights into the evolution of the chromosphere during flares. As the chromosphere mediates a large portion of the energy liberated during a flare, understanding its response is key to developing our understanding of the complete flaring system. It is important to verify the soundness of assumptions made in current modelling, and investigate effects that will become important as the next generation of high-resolution solar telescopes enter service. The work presented in this thesis seeks to develop our understanding of the formation and interpretation of spectral features in flares, which is a necessary step in determining the dominant energy transport mechanism in flares, as well as the associated atmospheric conditions.

In Chap. 2, we present the numerical approaches employed in flare modelling. Then, in Chap. 3 we describe the primary optical spectral lines that will be modelled, the instrumentation used, an overview of the inverse problem of radiative transfer, and an introduction to machine learning. The *Lightweaver* radiative transfer framework is presented and validated in Chap. 4. *Lightweaver* is then applied, in Chap. 5, to synthesising the radiation from current solar flare models to assess assumptions and investigate future directions. Chap. 6 presents the modelled radiative response of a quiet Sun slab adjacent to a flare. In Chap. 7, we describe our deep learning inversion technique, RADYNVERSION. Finally, in Chap. 8 concluding remarks are made.

### 1.3 Conventions

Throughout this thesis we use SI units unless otherwise noted, with wavelengths in nm. In all discussions the “height” or altitude of a point is considered to increase vertically from the solar surface (i.e. the layer of the photosphere which is opaque to light with a wavelength of 500 nm). This leads to a definition of velocity where a positive value indicates an observed blueshift.

When referring to different computers, we refer to:

1. *hercules*: 36 cores/72 threads across 2 sockets containing Intel Xeon E5-2697v4, 256 GB of DDR4 RAM, CentOS 7 Linux.

### 1.3 Conventions

2. *tomahna*: 8 cores/16 threads AMD Ryzen 3700X, 32 GB DDR4 RAM, Windows 10 with Ubuntu Windows Subsystem for Linux.
3. *hephaistos*: 6 cores/12 threads Intel i7 8700, 16 GB DDR4 RAM, CentOS 7 Linux.

This thesis is produced using PythonTeX (Poore 2015), and all figures are produced by Python code stored inline. Diagrams are produced using the Paul Tol’s vibrant colourblind accessible colour scheme<sup>2</sup>. The complete source of this document, and the data necessary to reproduce it, will be distributed and archived following typical practices. Whilst a complete list of packages will be made available with this distribution, the current version has been compiled with Python 3.8.2, NumPy 1.20.0 (Harris et al. 2020), SciPy 1.6.0 (Virtanen et al. 2020), Matplotlib 3.4.3 (Hunter 2007), seaborn 0.10.0 for its colourblind accessible colour scheme (Waskom 2021), Astropy 4.2.1 (Robitaille et al. 2013; Price-Whelan et al. 2018), and Lightweaver 0.7.5 (Osborne & Milić 2021). For all validations of the *Lightweaver* framework presented herein the data is produced using simulations run during the compilation of the document, and for the other chapters the complete post-processed simulation output or observational product is ingested, with the exception of Chap. 6 where the small region isolated in the observational data was extracted in a separate step to reduce the necessary volume of information. The source of this thesis therefore documents how this data was processed to produce the figures presented and can easily be modified to analyse the available data in other ways.

---

<sup>2</sup><https://personal.sron.nl/~pault/>

# 2

## Numerical Flare Modelling

Simulation is a powerful scientific approach that seeks both to validate our understanding of a phenomenon and also learn how it is influenced by various parameters. Contrary to its dictionary definition, which suggests deceit or merely a surface level resemblance, simulations are essential tools in an astrophysical context. Astrophysical observations differ greatly from laboratory experiments, due to our lack of knowledge of the configuration and parameters of the system, and inability to repeat particular events. Modelling therefore represents our best tool for bridging the gap between a theory of the physical processes at work in an observed event, and the observations thereof. A numerical approach is often needed as the coupled physical processes investigated are typically too complex and non-linear to analytically solve in detail. There will commonly be free parameters left in these models, some of which can be directly inferred from observation, but others may need to be investigated and later constrained by comparing the model output to observations. Due to computational or conceptual limitations, simplifying assumptions are often required to make these models tractable.

To design a model of solar flares it is first necessary to understand as much as possible from the available observations, so as to ascertain reasonable approximations and the most important processes to include. By the same, it is necessary to understand which observables vary between flares, to be able to exploit their specific sensitivities and choose those that exhibit sufficient variations to describe the processes at work and distinguish between models. The following builds on many years of work by the solar physics community, and we start from the understanding that has slowly been accrued to present the current state of flare modelling, and an introduction to some of the important numerical theory. A discussion of the flare observables we are focused on, as well as the techniques used to obtain these, is presented in Chap. 3.

## 2.1 Radiation Hydrodynamics

A complete numerical description of a solar flare would require considering the simulation of a large volume of plasma permeated by a strong magnetic field, using the equations of magnetohydrodynamics (MHD) to describe the motion of the fluid, and considering the propagation of energy through the volume. Due to the complexity of this treatment, including obtaining sufficient numerical resolution throughout the model, this is not currently feasible. It is reasonable to assume that the strength of the magnetic field effectively locks the fluid within flux tubes to the field lines. This is then used to construct an approach known as field-aligned radiation hydrodynamic (RHD) modelling. RHD modelling of flares starts from the assumption that the plasma to be modelled can be represented as being contained within a tube, often considered semi-circular. This represents a single magnetic flux tube, and the plasma may only move longitudinally along it. The atmosphere is then treated as a plasma (ranging from semi-ionised in the lower atmosphere to fully ionised in the corona) consisting of a single compressible fluid, obeying the equations of hydrodynamics and energy transport through this medium.

This gas dynamic system can be described by

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial z} &= 0, \\ \frac{\partial \rho v}{\partial t} + \frac{\partial \rho v^2}{\partial z} - \rho g + \frac{\partial P}{\partial z} - \frac{\partial}{\partial z} \left( \mu \frac{\partial v}{\partial z} \right) &= 0, \\ \frac{\partial E}{\partial t} + \frac{\partial}{\partial z} \left( E v + P v - \kappa \frac{\partial T}{\partial z} \right) + \mu \left( \frac{\partial v}{\partial z} \right)^2 - \rho v g + L - S &= 0, \end{aligned} \quad (2.1)$$

where  $\rho$  is the mass density of the gas,  $v$  is the bulk velocity,  $z$  is the spatial coordinate,  $g$  is the local gravitational acceleration,  $P$  is the pressure,  $\mu$  is the coefficient of dynamic viscosity,  $E$  is the total plasma energy,  $L$  is the radiative loss term, and  $S$  is all additional energy source terms, from radiative transfer and atmospheric heating. In these models the plasma is typically assumed to behave as an ideal gas, and therefore uses the equation of state  $P = n_{\text{tot}} k_B T$ , where  $n_{\text{tot}}$  is the total particle number density,  $k_B$  is Boltzmann's constant, and  $T$  is the plasma temperature. Additionally,  $\kappa$  is the temperature-dependent coefficient of heat conduction, and is discussed in detail in Sec. 2.6.

All of the quasi-one dimensional codes solve a variant of the RHD equations presented in (2.1). The first of these describes mass continuity, the second conservation of momentum, and the third conservation of energy. The nuance typically occurs in ensuring that all

## 2.1 Radiation Hydrodynamics

terms are solved in a self-consistent fashion and the treatment of the optically thick radiation source and sink terms appearing in the energy equation.

The first generation of hydrodynamic flare models (Nagai 1980; Mariska et al. 1982; McClymont & Canfield 1983) focused on capturing the gas dynamics of the flaring event, using simplified radiative treatments. These early models were, in general, much more limited due to computational constraints. Nagai (1980) and Mariska et al. (1982) both employ optically thin losses with ad hoc corrections to avoid over-cooling the lower atmosphere. McClymont & Canfield (1983) employed an escape probability formalism for determining the transition rates of hydrogen, which is a faster, less precise method than detailed radiative transfer, but vastly more accurate than assuming the chromosphere to be optically thin (or an ad hoc modulation thereof). Escape probability methods are unable to directly take radiative backwarming into account – where radiation produced higher in the atmosphere is absorbed at a lower point – which can be an important factor for strong solar optical transitions used to diagnose atmospheric properties.

The current generation of RHD codes includes RADYN (Carlsson & Stein 1992, 1995, 1999; Allred et al. 2015), FLARIX (Varady et al. 2010; Heinzl et al. 2015), HYDRAD (Bradshaw & Mason 2003; Bradshaw & Cargill 2013), and HYDRO2GEN (Druett & Zharkova 2018, 2019). RADYN and FLARIX both apply detailed treatment of the radiative losses for certain chromospheric transitions, whilst HYDRAD uses partially precomputed radiative rates for hydrogen (based on the treatment of Sollum (1999)) and radiative losses following the approximations of Carlsson & Leenaarts (2012). HYDRAD originally focused more on the investigation of non-equilibrium optically thin radiative losses in the corona, and less on the chromosphere, however, recent development has also moved in the direction of an improved chromospheric treatment. It also differs in its use of a two-fluid model, treating electrons and ions as separate, but coupled, fluids. In current applications this does not appear to provide significantly different results when applied to flare modelling. HYDRO2GEN applies a different approach to computing the radiative transfer terms using the approximate L2 method of Ivanov & Serbin (1984), and electron beam heating following Syrovatskii & Shmeleva (1972).

RADYN is derived from the MULTI radiative transfer code (Scharmer & Carlsson 1985; Carlsson 1986, 1992), and applies the linearisation approach described therein to the entire RHD system, ensuring the self-consistency of all terms. The equations of RHD are solved in an implicit linearised form by Newton-Raphson iteration on the dynamic grid of Dorfi & Drury (1987). RADYN was originally constructed to investigate the effects of waves propagating through the chromosphere, but was extended by Abbett & Hawley

## 2.2 An Eye to the Future: Radiative Magnetohydrodynamics

(1999) to perform some of the earliest simulations of the dynamic flares with detailed radiation treatment in a self-consistent fashion. The code has since been developed further to include improved energy transport treatments, primarily focusing on electron beam energy deposition (Allred et al. 2005, 2015). Its techniques are discussed further in Sec. 5.1.

FLARIX, on the other hand, consists primarily of three separate modules. The hydrodynamics is solved by a variant of the NRL Solar Flux Tube Model described in Mariska et al. (1982, 1989), extended with a MALI module (considering non-overlapping spectral lines with constant background continuum as described in Rybicki & Hummer (1991)) for detailed radiation transfer, and a test-particle code for determining heating and electron deposition (Varady et al. 2010; Heinzel et al. 2015).

RADYN and FLARIX have recently been tested against each other and provide remarkable agreement, despite their different heritage (Kašparová et al. 2019), and early tests also show good consistency with HYDRAD when its recent extensions are enabled<sup>1</sup>.

## 2.2 An Eye to the Future: Radiative Magnetohydrodynamics

Much scientific effort is being invested into the creation of three-dimensional radiative MHD (RMHD) models, which aim to accurately describe all of the physics of a flaring system through the use of large models on supercomputers. The two leading codes are *Bifrost* (Gudiksen et al. 2011), and MURaM (Rempel et al. 2009; Rempel 2016). MURaM was originally focused on photospheric magnetoconvection, but has been extended into the corona. It adopts a grey radiative transfer technique that treats all outgoing radiation using a wavelength averaged approach, but has successfully demonstrated the ability to drive self-consistent flare-like eruptions with time-dependent photospheric magnetograms as a boundary condition (Cheung et al. 2019).

*Bifrost* is more focused on a detailed chromospheric treatment; it can apply different combinations of physics modules including multi-group opacities for wavelength dependent radiation transport, the Sollum (1999) model for hydrogen ionisation (following the treatment of Leenaarts et al. 2007), and radiative losses following the method of Carlsson & Leenaarts (2012). It has been used to investigate solar enhanced networks (Carlsson et al. 2016), the generation of spicules (Martínez-Sykora et al. 2017), and the magnitude

---

<sup>1</sup>Initial comparisons undertaken following the International Space Science Institute meeting: “Interrogating Field-Aligned Solar Flare Models: Comparing, Contrasting and Improving” led by G.S. Kerr and V. Polito.

## 2.3 Solar Flare Heating

of magnetic energy present in the chromosphere (Martínez-Sykora et al. 2019), amongst many other projects.

Currently, neither code can handle the scale of the energy deposition used in the field-aligned flare models, or provide an equivalent chromospheric spatial resolution, but this is an area of very active development. The field-aligned models remain complementary to these much more complex and computationally intensive models, allowing for more rapid investigation of the relative importance of phenomena to be implemented within the RMHD models, the use wider parameter spaces and higher energy inputs, and easier more compact output that is easier to manipulate and interpret.

### 2.3 Solar Flare Heating

The aforementioned modern field-aligned RHD codes are primarily designed to simulate the response of a tube of initially quiet solar atmosphere to heating. This is typically presumed to be due to a beam of energetic electrons precipitating from the corona due to magnetic reconnection.

The direct effects of these electrons must also be considered, as their flux is sufficiently large that their collisions with particles in the lower atmosphere can substantially change the distribution of these particles, exciting and ionising them. The evolution of the atmosphere will also affect how and where it interacts with these precipitating electrons, coupling an additional problem to the extant RHD system.

There are different methods of solving this problem, the simplest of which is the analytic solution of Emslie (1978), known as the “Emslie beam”, which provides an analytic solution to electron beam energy deposition along the column of plasma under the assumption that all electron energy is lost due to Coulomb collisions with the plasma acting as a cold target. This is a useful approximate treatment, however terms such as relativistic effects on high energy electrons, return current (heating the corona), magnetic mirroring, and particle diffusion (pitch-angle and momentum) are all important here, and ignored in this model. RADYN and FLARIX both include more advanced treatments including these effects; FLARIX uses a test-particle module (Varady et al. 2010), whereas RADYN solves the Fokker-Planck equation (now using the method of Allred et al. 2020), but both also include the option to use the the simpler analytic Emslie formulation.

## 2.4 Introduction to Radiative Transfer

The Alfvénic heating model has also been considered in modelling efforts. A simplified variant of this has been investigated by [Kerr et al. \(2016\)](#) inside the RADYN code, and produced notably different spectral line profiles which may agree better with observations than those from an electron beam model. Further investigation of a more complete implementation of this method and its results are needed. This approach was also incorporated into HYDRAD by [Reep & Russell \(2016\)](#), which showed the viability of these waves heating the deep chromosphere and triggering explosive chromospheric evaporation. It is non-trivial to incorporate a full treatment of this phenomenon into RHD codes as it is coupled to the magnetic field, and future improvements are needed to more accurately model these effects and investigate to what extent Alfvénic and electron beam heating occurs simultaneously.

The simulations of flares discussed thus far allow us to model the time-evolution of heating a starting atmosphere and in turn predict line and continuum emission through the tools of radiative transfer. These detailed results can be compared against observations, and have provided significant insight into chromospheric properties (e.g. [Kuridze et al. \(2015\)](#); [Rubio da Costa et al. \(2016\)](#); [Kowalski et al. \(2017b\)](#); [Simões et al. \(2017\)](#)). The manual forward-fitting process involved in attempting to reproduce these observations with simulations is both time-consuming and difficult and lies close to the field of automated inversions, which we will address later in Chaps. 3 and 7.

We will now present a more detailed overview of the physics involved in the most important terms of the RHD equations (2.1), starting with complexities of radiative transfer outside of local thermodynamic equilibrium.

## 2.4 Introduction to Radiative Transfer

*The content of this section draws primarily from [Hubeny & Mihalas \(2014\)](#) and the paper describing the *Lightweaver* radiative transfer framework ([Osborne & Milić 2021](#)).*

Radiative transfer is the science of how radiation propagates through a material: the absorptions, scatterings and emission that happen therein. Radiation is by far the most widely exploited conduit through which information can arrive from celestial bodies. It allows us to derive proxies for *in situ* measurements, that cannot otherwise be obtained, due to the distances and extreme conditions being observed.

## 2.4 Introduction to Radiative Transfer

Everyone is familiar with the concept of images, which show the spatial variation of light, but a lot of additional information can be gleaned by analysing radiation in terms of its wavelength variation and polarisation projections. Using both of these properties is referred to as *spectropolarimetry*, and *spectroscopy* when only the unpolarised intensity is considered. The most basic property to consider here is the specific intensity, commonly denoted  $I(\nu, \vec{d})$  for a particular frequency  $\nu$  and direction  $\vec{d}$  at a location in space with typical SI units  $\text{J m}^{-2} \text{s}^{-1} \text{Hz}^{-1} \text{sr}^{-1}$ .

A ray propagating through a medium, such as a neutral gas or plasma, will gain a certain amount of energy per unit length due to emission processes, whilst also losing another amount due to absorption and scattering processes. These will depend on the local parameters of the plasma as well as the direction of the ray. For a plasma where the primary interacting species are atomic, we can distinguish bound-bound (spectral lines) and bound-free (continuum) transitions. In the former a bound electron moves between two different sublevels of an atom<sup>2</sup>, whilst in the latter the atom either absorbs sufficient energy to free a previously bound electron, or a free electron recombines with an atom and loses energy in the process.

In addition to the obvious spontaneous emission and absorption processes, there is also a process of stimulated emission which is needed to balance the transitions. This occurs when an electron is stimulated to transition between levels by photons with the same frequency and direction as the photon produced by this transition.

In the following, we will discuss how to obtain the frequency- and direction-dependent outgoing radiation given the emissivity and opacity of the plasma, and how to obtain a self-consistent radiation field and populations for atomic species outside of the approximations of local thermodynamic equilibrium. We will also discuss how to determine convergence for the iterative processes used.

### 2.4.1 The Formal Solution

For a one-dimensional planar atmosphere the radiative transfer equation (RTE) for the specific intensity along a ray can be written as

$$\frac{1}{c} \frac{\partial I(\nu, \vec{d})}{\partial t} + \mu \frac{\partial I(\nu, \vec{d})}{\partial z} = \eta(\nu, \vec{d}) - \chi(\nu, \vec{d})I(\nu, \vec{d}), \quad (2.2)$$

---

<sup>2</sup>Here *atom* refers to either a neutral atom or an ion

## 2.4 Introduction to Radiative Transfer

where  $\eta$  is the plasma emissivity,  $\chi$  is the plasma opacity,  $z$  is the spatial coordinate of the stratification of the atmosphere,  $t$  represents time,  $c$  is the speed of light, and  $\mu$  is the cosine of the ray's inclination to the surface normal of the plane-parallel configuration. We consider that the light-crossing time for the propagation of light on a solar scale is small compared to the time evolution of both the atmosphere and our observations so we ignore the time-derivative term. Defining the source function

$$S(\nu, \vec{d}) = \frac{\eta(\nu, \vec{d})}{\chi(\nu, \vec{d})}, \quad (2.3)$$

and the optical depth along a ray from the observer as the number of photon mean free paths along this segment

$$\tau(z, \nu, \vec{d}) = \int_z^{z_{\text{obs}}} \frac{\chi(z', \nu, \vec{d})}{\mu} dz', \quad (2.4)$$

we can write the RTE as

$$\frac{\partial I(\nu, \vec{d})}{\partial \tau(\nu, \vec{d})} = I(\nu, \vec{d}) - S(\nu, \vec{d}). \quad (2.5)$$

Equation (2.5) is a first-order linear differential equation and can be solved with the integrating factor  $e^{-\tau(\nu, \vec{d})}$  giving the formal solution of the RTE:

$$I(\tau_0, \nu, \vec{d}) = I(\tau_1, \nu, \vec{d})e^{-(\tau_1 - \tau_0)} + \int_{\tau_0}^{\tau_1} S(t_\nu, \nu, \vec{d})e^{-(t_\nu - \tau_0)} dt_\nu, \quad (2.6)$$

for  $\tau_0$  the optical depth at the observer,  $t_\nu$  a dummy variable used for integration, and  $\tau_1 > \tau_0$  along the line of sight.

The solution in equation (2.6) prescribes nothing about the form of the source function in the atmosphere, and assumes that it varies continuously. A *formal solver* is a numerical routine that computes the intensity in a discretised atmosphere by solving a form of the RTE (2.5). There are approaches such as that of [Feautrier \(1964\)](#) that involve casting the problem as a second order differential equation for both an up-going and a down-going rays simultaneously, it cannot handle both Doppler shifts and overlapping lines, and for these reasons we will not discuss it further. We shall instead focus on the so-called short-characteristic method consisting of solving the RTE directly between discrete points by prescribing a functional form for its variation between these points.

### 2.4.2 Short-Characteristics Methods

If a functional form with an analytic integral is chosen for the variation of the source function between defined points then an atmosphere can be treated as a sum of analytic integrals. We shall consider the simplest useful functional form, a linear variation, as an illustrative example. This approach was first presented by [Olson & Kunasz \(1987\)](#).

The RTE for one frequency and direction through a slab of a plane-parallel atmosphere (in the case of outgoing radiation (i.e.  $\mu > 0$ )) can be written

$$I(\tau_0) = I(\tau_1) \exp(-|\tau_0 - \tau_1|) + \int_{\tau_0}^{\tau_1} S(t) \exp(-(t - \tau_0)) dt, \quad (2.7)$$

with  $t$  as a dummy integration variable.

Now, assuming a linear variation of  $S$  with  $\tau$  in this slab gives

$$S(t) = S_{\tau_0} \frac{\tau_1 - t}{\tau_1 - \tau_0} + S_{\tau_1} \frac{t - \tau_0}{\tau_1 - \tau_0}, \quad (2.8)$$

where  $S_{\tau_i}$  indicates the value of the source function at optical depth  $\tau_i$ . This can then be substituted into (2.7) (with  $\Delta := \tau_1 - \tau_0$ ) giving

$$I(\tau_0) = I(\tau_1) \exp(-|\tau_1 - \tau_0|) + \frac{\Delta - 1 + \exp(-\Delta)}{\Delta} S_{\tau_0} + \frac{1 - \exp(-\Delta) - \Delta \exp(-\Delta)}{\Delta} S_{\tau_1}. \quad (2.9)$$

This expression can be applied repeatedly from one end of the atmospheric model (assuming that the incoming radiation field is known), to the other, to provide the emergent intensity. It is this procedure we refer to as computing a formal solution, and this will typically need to be computed for multiple directions (angles to the surface normal in the plane-parallel case) to compute the angle-averaged radiation field at each frequency and location in a model atmosphere using a weighted quadrature, such as the Gauss-Legendre quadrature.

In atmospheres with very well resolved spatial grids, this method works quite well, however whenever the true source function has positive curvature, the intensity is overestimated, and underestimated for negative curvature. These effects can become quite significant in more sparsely sampled atmospheres.

The short characteristics method can be improved by using higher order polynomial interpolants, however these can lead to spurious ringing artifacts, negatively affecting their precision. One commonly used robust formulation is the monotonic piecewise

## 2.4 Introduction to Radiative Transfer

parabolic method of [Auer & Paletou \(1994\)](#). This method assumes a parabolic variation of the source function across three consecutive points (i.e. a three point stencil), but limits it to the value obtained from linear interpolation if the parabolic interpolant exits the range bounded by these three points.

Other interpolating functions can be used. For example, the cubic Bézier spline technique of [de la Cruz Rodríguez & Piskunov \(2013\)](#), provides a higher order approximation in regions of smooth variation, and can be limited through the control points to prevent any ringing instability. A similar approach has been taken with the BESSER quadratic Bézier spline approach of [Štěpán & Trujillo Bueno \(2013\)](#). For the plane-parallel case, all of these methods can be derived analogously to the linear formal solver presented above. In [Sec. 6.1](#) we describe the implementation of the BESSER method in a two-dimensional atmosphere.

### 2.4.3 Other Formal Solvers

[Janett et al. \(2018\)](#) propose a novel approach to the formal solver, using an optimised solver for the differing optical thickness in each slab. They show that this leads to substantial performance benefits whilst also being more numerically stable. They comment that whilst higher order formal solvers will theoretically converge better to the true result, due to the assumptions that are made in their derivation, this will only occur if the variation of the source function in relevant regions of the atmosphere is sufficiently smooth. The modern three-dimensional RMHD simulations use relatively coarse spatial grids with large transients in atmospheric parameters that risk provoking instability in the higher order formal solvers, especially in the case of full Stokes radiative transfer.

As work continues on these higher resolution RMHD simulations an investigation into how each formal solver handles discontinuous parameters is needed. This has been discussed by [Steiner et al. \(2016\)](#), and the methods employed for reconstructing discontinuous parameters in numerical hydrodynamics (see [Sec. 2.5.3](#) for a description of these) may present a future avenue for more robust, accurate, and efficient formal solvers. An accurate treatment of steep gradients and discontinuities is not just valuable to the RMHD models, but important to all radiative transfer modelling, as a reduction in the spatial resolution needed to correctly evaluate radiative transfer can lead to significant reductions in computational requirements, making more complex simulations possible.

#### 2.4.4 LTE vs NLTE

With the understanding of how to design and implement formal solvers for the radiative transfer equation, we have the ability to compute the radiation leaving an atmosphere from the opacity and emissivity structure. The values of these depend on the atomic populations, and their distribution across energy levels and ionisation states.

As stars like the Sun radiate energy into the cosmos, it is clear that they cannot be in total thermal equilibrium. However, it is reasonable to suggest that if the atmosphere is sufficiently collisional then a form of local thermodynamic equilibrium (LTE) holds, whereby local parcels of the plasma are effectively in thermodynamic equilibrium such that Kirchhoff's laws of radiation hold. When a species within a plasma is in LTE its population distribution can be computed using the Saha-Boltzmann equation. This is the case for spectral lines that form in the photosphere, and the associated level populations, emissivity, and opacity can all be computed directly from local thermodynamic quantities<sup>3</sup>.

If the radiative rates instead dominate the total transitional rates for the species then the particle distribution can no longer be governed by purely local parameters and we enter the realm of non-LTE (NLTE) physics. This occurs as we enter the chromosphere, where the collisional rates decrease, and the radiation field couples the atomic populations in different regions together. The core focus of NLTE radiative transfer is to determine atomic populations consistent with the local thermodynamic parameters of the atmosphere and the non-local radiation field. Until otherwise specified, we will consider that the electron density throughout the atmosphere is known and consider it part of the atmospheric model.

Expressing this mathematically, we write the total transition rate  $P_{ij}$  per atom in level  $i$  between levels  $i$  and  $j$  (with the convention that  $i < j$ ) as

$$P_{ij} = R_{ij} + C_{ij}, \quad (2.10)$$

where  $R_{ij}$  is the rate of radiative transitions (due to absorption, spontaneous, and stimulated emission) and  $C_{ij}$  is the rate of collisional transitions. As the total population of each element must remain constant we can write the kinetic equilibrium equation (which

---

<sup>3</sup>If the electron density is not known *a priori* then an iteration scheme using the Saha-Boltzmann equation is necessary to determine consistent values of both the electron density and the atomic populations.

## 2.4 Introduction to Radiative Transfer

can be derived from the Boltzmann equation)

$$\frac{\partial n_l}{\partial t} + \nabla \cdot (n_l \vec{v}_s) = \sum_{l' \neq l} (n_{l'} P_{l'l}) - n_l \sum_{l' \neq l} P_{ll'}, \quad (2.11)$$

where  $n_l$  is the number density of atoms in level  $l$  of the atomic species in question and  $\vec{v}_s$  is the bulk velocity of this species.

Equation (2.11) is frequently simplified to the statistical equilibrium equation, by setting the left-hand side to 0 giving

$$\sum_{l' \neq l} (n_{l'} P_{l'l}) - n_l \sum_{l' \neq l} P_{ll'} = 0. \quad (2.12)$$

In Chaps. 5 and 6 we will discuss when the full time-dependent solution need be considered over the simpler statistical equilibrium solution. The former of these requires a history of the atomic populations in the atmosphere, typically from a time-evolving model such as an RHD simulation, whereas the latter associates a unique solution of atomic populations to a given atmosphere.<sup>4</sup> In the following, even when considering the time-dependent problem, we assume that the timescale over which each step in our numerical simulation is integrated is long compared to the light-crossing time for the regions where each transition is optically thick. If this is not the case, then a different formulation of the formal solver will be required, so as to solve (2.2).

### 2.4.5 Collisional Rates

There are many collisional processes by which electrons can transition between energy levels in a plasma, including ionisation and recombination processes. These include excitation of ions by electrons, ionisation and excitation of neutral by electrons, excitation by protons and neutral hydrogen, as well as charge exchange with these species. There are more advanced rate formulations that depend on complex functional forms derived from laboratory work and theoretical analysis, such as those of [Burgess & Chidichimo \(1983\)](#); [Arnaud & Rothenflug \(1985\)](#).

In most cases, the collisional rates are assumed to only depend on local plasma properties, such as the temperature, total particle density, and electron density, with the particles

---

<sup>4</sup>The uniqueness of the solution was proven by [Rybicki \(1997\)](#) for the linear case of (2.12). I have found no equivalent proof for the non-linear system when a variable electron density is also considered.

locally following Maxwellian velocity distributions that connect the upwards and downwards collisional rates for a process through detailed balance i.e.,  $n_i^* C_{ij} = n_j^* C_{ji}$ , where  $n_i^*$  indicates the LTE population of level  $i$ . This is not the case for non-thermal rates, such as those of Fang et al. (1993) which consider collisional excitation by non-thermal electrons from the electron beams used in RHD simulations, but are still computed from local parameters.

### 2.4.6 Emissivity and Opacity

To mathematically formulate the radiative rates that are so important to  $P_{ij}$  in a NLTE context we must first formulate expressions for emissivity and opacity. There are two forms of transition we need to consider for atomic radiative transfer, spectral lines, and continua. In spectral lines, the three processes that need to be understood to develop a model for emissivity and opacity are spontaneous and stimulated emission, and absorption. The magnitude of these terms are controlled by the Einstein coefficients  $A_{ji}$ ,  $B_{ij}$ , and  $B_{ji}$  respectively, which are related by the Einstein relations

$$g_i B_{ij} = g_j B_{ji}, \quad (2.13)$$

where  $g_i$  is the statistical weight of level  $i$  and

$$A_{ji} = \frac{2h\nu_{ij}^3}{c^2} B_{ji}, \quad (2.14)$$

where  $h$  is Planck's constant,  $c$  is the speed of light, and  $\nu_{ij}$  is the rest frequency of the spectral line.

The rest frequency  $\nu_{ij}$  of a spectral line is defined by

$$\nu_{ij} = \frac{\Delta E_{ji}}{h}, \quad (2.15)$$

where  $\Delta E_{ji}$  is the energy difference between levels  $j$  and  $i$ . Bound-bound transitions between states in a static plasma are not infinitely narrow, but are instead broadened by a number of factors such as natural broadening from uncertainty in the lifetime of the upper state, Doppler broadening due to random thermal motions in the plasma, and collisional broadening (e.g. van der Waals, and Stark broadening<sup>5</sup>). The net effect of these processes

---

<sup>5</sup>van der Waals broadening is due to an interaction between an excited atom and the dipole it induces over a neutral atom, whilst Stark broadening is due to the interaction between an atom and a charged particle.

## 2.4 Introduction to Radiative Transfer

typically leads to spectral line profiles being modelled as a Voigt function (the convolution of a Gaussian and a Lorentzian). The Gaussian terms are due to Doppler broadening, whilst the other terms are typically modelled as Lorentzians (e.g. Sutton 1978), although more accurate treatments of Stark broadening such as those employed by Kowalski et al. (2017b) are non-Lorentzian, and must be separately convolved with the Voigt profile produced from the previous terms. This term typically produces chromospheric lines much broader than those resulting from the method of Sutton (1978), and becomes important as the density of charged particles in the chromosphere increases during flares, or in earlier spectral type (i.e. hotter) stars. The normalised line absorption profile then describes the probability of a photon with a certain energy being absorbed by the transition.

Typically it is assumed that the plasma is sufficiently collisional for elastic collisions to redistribute the electrons of an atom in an excited state over all sub-states of an energy level prior to emission. If this is not the case then the frequency of the outgoing photon will be correlated with the frequency of the photon that excited the atom into this state. Thus the spectral line will have an emission profile distinct from its absorption profile. This effect is known as partial frequency redistribution (PRD) and will be discussed in Sec. 2.4.12. The typical state of affairs, where there is sufficient redistribution from elastic collisions for these two processes to be uncorrelated is known as complete redistribution (CRD). In the following, we will always express line emission processes through an emission profile for generality, even if they are treated as CRD.

Continua, or bound-free transitions, depend instead on photoionisation cross-sections. In the case of hydrogenic ions these fall off with  $1/\nu^3$  (for photons of frequency  $\nu$ ) as the energy of the ionising photon increases away from the continuum edge. The edge of a continuum is defined by the ionisation potential for an atom in the bound state of this transition, as only photons with an energy equal to or greater than this are capable of photoionising the element. In general these cross-sections and their variations are determined from laboratory experiments and numerical solutions of the Schrödinger equation. The photoionisation cross-section describes how likely the atom is to interact with a photon of a particular energy through this bound-free transition. Its partner processes (like spontaneous and stimulated emission for bound-bound transitions) are spontaneous and stimulated recombination, whereby an electron is captured by the atom and a photon is released with energy equal to the excess energy of the system. These are related to the photoionisation cross-section and the local plasma parameters by the Milne relations.

There are also free-free interactions between particles (often known as *bremstrahlung*),

## 2.4 Introduction to Radiative Transfer

where electrons and ions interact, and the electron gains or loses energy through absorption or emission of a photon. This process also has a characteristic cross-section and is considered as part of the background opacities and emissivities.

Following the notation of Rybicki & Hummer (1992) and Uitenbroek (2001) the emissivity  $\eta$  and opacity  $\chi$  for a transition can then written

$$\eta_{ij} = n_j U_{ji}(\nu, \vec{d}), \quad (2.16)$$

$$\chi_{ij} = n_i V_{ij}(\nu, \vec{d}) - n_j V_{ji}(\nu, \vec{d}). \quad (2.17)$$

The U and V terms are defined for bound-bound and bound-free transitions as

$$U_{ji} = \begin{cases} \frac{h\nu}{4\pi} A_{ji} \psi_{ij}(\nu, \vec{d}), & \text{bound-bound} \\ n_e \Phi_{ij}(T) \left( \frac{2h\nu^3}{c^2} \right) e^{-h\nu/k_B T} \alpha_{ij}(\nu), & \text{bound-free,} \end{cases} \quad (2.18)$$

$$V_{ij} = \begin{cases} \frac{h\nu}{4\pi} B_{ij} \phi_{ij}(\nu, \vec{d}), & \text{bound-bound} \\ n_e \Phi_{ij}(T) e^{-h\nu/k_B T} \alpha_{ij}(\nu), & \text{bound-free,} \end{cases} \quad (2.19)$$

$$V_{ji} = \begin{cases} \frac{h\nu}{4\pi} B_{ji} \psi_{ij}(\nu, \vec{d}), & \text{bound-bound} \\ \alpha_{ij}(\nu), & \text{bound-free,} \end{cases} \quad (2.20)$$

where  $\phi$  is the line absorption profile,  $\psi$  is the line emission profile,  $\alpha_{ij}$  is the photoionisation cross-section,  $n_e$  is the local electron density, and  $k_B$  is the Boltzmann constant.  $\Phi$  is the Saha-Boltzmann equation given by

$$n_e \Phi_{ij}(T) = \frac{n_i^*}{n_j^*} = \frac{g_i}{2g_j} \left( \frac{h^2}{2\pi m_e k_B T} \right)^{3/2} \exp\left( \frac{\Delta E_{ji}}{k_B T} \right), \quad (2.21)$$

where  $n^*$  is the population of the species in LTE,  $m_e$  is the electron mass,  $\Delta E_{ji}$  is the energy difference between levels  $j$  and  $i$ , and  $g_i$  is the statistical weight of level  $i$ . By convention we define  $U_{ij} = U_{ii} = V_{ii} = 0$  and  $\chi_{ij} = -\chi_{ji}$ . From these definitions we see that the U quantities relate to spontaneous emission, whereas the V quantities describe stimulated processes. In the case of complete redistribution, where  $\psi = \phi$ , all of the U and V terms are constant at each location in a given atmosphere, and the variation in emissivity and opacity depends on the atomic populations.

### 2.4.7 Radiative Rates

The radiative rates that are needed to solve the kinetic or statistical equilibrium equations are an expression of the number of upward (ij) and downward (ji) transitions due to absorption, spontaneous and stimulated emission processes (and the equivalent bound-free processes). The formulation of emissivity and opacity through  $U$  and  $V$  also allows us to express the radiative rates succinctly for both upwards and downwards transitions in lines and continua at each location in the atmosphere as

$$R_{l\nu} = \oint \frac{1}{h\nu} \left( U_{l\nu}(\nu, \vec{d}) + V_{l\nu}(\nu, \vec{d}) I(\nu, \vec{d}) \right) d\nu d\Omega, \quad (2.22)$$

where  $I(\nu, \vec{d})$  is the specific intensity at this location for a given frequency  $\nu$  and direction  $\vec{d}$ . It is through  $I$  that the non-locality of the radiation field enters the problem.

### 2.4.8 General Source Function

Where multiple atomic species are present, the total emissivity and opacity are simply the sum of the emissivity and opacity for every transition on each atom at the current frequency and direction. It is common to additionally consider scattering by processes such as Thomson scattering, in which case the source function will be written

$$S(\nu, \vec{d}) = \frac{\eta_{\text{tot}}(\nu, \vec{d}) + \sigma(\nu)J(\nu)}{\chi_{\text{tot}}(\nu, \vec{d})}, \quad (2.23)$$

where the “tot” subscript refers to these terms being summed over all interacting species,  $\sigma$  describes the frequency-dependent coherent and isotropic continuum scattering cross-section, and

$$J(\nu) = \frac{1}{4\pi} \oint I(\nu, \vec{d}) d\Omega \quad (2.24)$$

is the angle-averaged intensity at frequency  $\nu$ .

### 2.4.9 Iterative Solutions

Now we have an expression for the radiative rates in each transition that can be computed numerically given the local value of the intensity, however the radiation field is not known *a priori*. Clearly, an iterative scheme will be needed to find a stable set of populations yielding a self-consistent radiation field.

## 2.4 Introduction to Radiative Transfer

If we treat the formal solver as an operator  $\Lambda$  yielding the intensity from the source function throughout the atmosphere, i.e.

$$I(\nu, \vec{d}) = \Lambda_{\nu, \vec{d}}[S(\nu, \vec{d})], \quad (2.25)$$

then starting from an initial estimate of the atomic populations (e.g. LTE) we can compute the radiation field throughout the atmosphere and iteratively use this to update the populations by solving (2.11). This is known as Lambda iteration (after the operator used) and presents woefully poor convergence in optically thick conditions as the size of the population updates stagnate long before the true NLTE populations are obtained. Each additional Lambda iteration performed effectively accounts for photons that were scattered an additional time (i.e. the first Lambda iteration accounts for photons unscattered after emission, the second for once-scattered...). In the cores of optically thick lines photons will scatter a vast number of times and thus an equivalent number of Lambda iterations will be required.

The failure of Lambda iteration can be remedied by a process known as operator splitting, first introduced by Cannon (1973) whereby we set

$$\Lambda = \Lambda^* + (\Lambda - \Lambda^*), \quad (2.26)$$

with  $\Lambda^*$  an approximation of  $\Lambda$ . The iterative scheme then becomes

$$I(\nu, \vec{d}) = \Lambda_{\nu, \vec{d}}^*[S(\nu, \vec{d})] + (\Lambda_{\nu, \vec{d}} - \Lambda_{\nu, \vec{d}}^*)[S^\dagger(\nu, \vec{d})], \quad (2.27)$$

where  $\dagger$  identifies values from the previous iteration. This method is termed accelerated Lambda iteration (ALI), and can be shown to accelerate convergence by significantly amplifying the size of the corrections that would be computed by Lambda iteration at large optical depths, for an appropriately chosen  $\Lambda^*$ . From (2.27) we can see that it is necessary to invert  $\Lambda^*$  to obtain the updated value of  $S$ , on which  $I$  is also dependent. For a two-level atom this is discussed at length in Chaps. 12 and 13 of Hubený & Mihalas (2014). The full coupling of the terms in the multilevel NLTE problem will be made explicit in Sec. 2.4.10 when the MALI methods are presented; for now it is clear that the source function depends on the emissivity and opacity of each species, which in turn are controlled by the atomic populations, which are affected by the radiation field.

A good choice for  $\Lambda^*$  is not immediately evident, as it should be cheap to construct and invert, whilst providing a good approximation of  $\Lambda$ . Scharmer (1981) presented an upper triangular approximate operator that fits these criteria and showed its relation to the core-

## 2.4 Introduction to Radiative Transfer

saturation approach of [Rybicki \(1972\)](#) (where the net rates in the line core and wing are treated separately to precondition the net radiative rates by removing the large proportion of photons that are emitted in the wing and immediately reabsorbed).

[Olson et al. \(1986\)](#) proposed the use of the diagonal of the true  $\Lambda$  operator as an approximate operator, and showed that this is close to optimal, and is clearly trivial to invert (as it is a scalar at each location in the atmosphere). Now, the diagonal of  $\Lambda$  is easy to obtain by setting a test source function  $S = \delta_{ad}$  (where  $\delta$  is the Kronecker delta) and computing

$$\Lambda^* = \Lambda[S]. \quad (2.28)$$

Taking the example of the linear short characteristic formal solver presented in [Sec. 2.4.2](#) and substituting this definition of  $S$  we obtain

$$\Lambda_{v,\bar{d}}^* = \frac{\Delta - 1 + \exp(-\Delta)}{\Delta}, \quad (2.29)$$

where  $\Delta$  is defined as in [Sec. 2.4.2](#). The approximate operator can be computed analogously for other formal solvers.

### 2.4.10 Solving the Multilevel NLTE Problem

Starting from the radiative transfer equation (2.5) and the kinetic equilibrium equation (2.11) we can construct a framework with which to solve the multilevel NLTE problem. The primary term of interest is the right-hand side of (2.11), which is concerned with the atomic transition rates. This is also shared with the statistical equilibrium equations (2.12), and thus we will solve the latter of these and return to the former later. We follow the approach of [Rybicki & Hummer \(1992\)](#) and [Uitenbroek \(2001\)](#), which is known as Multi-level Accelerated Lambda Iteration (MALI) with full-preconditioning. The full-preconditioning approach handles arbitrary overlaps of lines and continua for multiple multilevel atoms. Earlier MALI methods ([Rybicki & Hummer 1991](#)) are capable of handling lines overlying a constant background continuum, but not interacting with each other. These methods describe a form of radiative transfer that is still based on a transition by transition approach to determining the radiative rates (which does not preclude self-consistency). [Socas-Navarro & Trujillo Bueno \(1997\)](#) proved that this approach can be numerically equivalent to the linearised ALI method of [Scharmer & Carlsson \(1985\)](#) implemented in the MULTI code ([Carlsson 1986, 1992](#)) provided the same assumptions are made in the design of the local operator. Note that we distinguish between the linearised

## 2.4 Introduction to Radiative Transfer

ALI method and the older method of complete linearisation (e.g. [Auer & Mihalas 1969](#); [Auer 1973](#); [Auer & Heasley 1976](#)). The latter of these is significantly more computationally demanding and less numerically stable than the ALI based methods we are considering here (a comparison of MULTI and LINEAR-B is shown in [Carlsson 1986](#)).

As MALI with full-preconditioning directly supports all forms of overlapping transitions, and the interactions between them, whilst retaining the convergence properties of the ALI method, we focus solely on this method in the following description of radiative transfer. Whilst there are also more rapidly converging methods (which can be viewed as extensions and variants of the ALI approach) such as the Gauss-Seidel and successive-over-relaxation methods of [Trujillo Bueno & Fabiani Bendicho \(1995\)](#) (with the extension to multilevel systems presented by [Paletou & Léger \(2007\)](#)), the forth-and-back implicit Lambda iteration of [Atanacković-Vukmanović et al. \(1997\)](#) (extended to multilevel systems by [Kuzmanovska et al. \(2017\)](#)), and even the multi-grid method of [Fabiani Bendicho et al. \(1997\)](#), we opt for this well-tested method that is proven stable and reliable in a wide variety of situations. This is particularly key when considering modelling of flares, which present much larger variations in parameters and steeper gradients than the quiet Sun models or academic semi-empirical models often considered. The other, more rapidly convergent, methods present possible avenues of improvement for *Lightweaver* in the future.

Now, following [Rybicki & Hummer \(1992\)](#) and [Uitenbroek \(2001\)](#) and substituting (2.25) into (2.12), and subsequently expanding the radiative rates gives

$$\begin{aligned} \sum_{l' \neq l} (n_{l'} C_{l'l}) + \sum_{l' \neq l} \oint \int \frac{1}{h\nu} n_{l'} (U_{l'l}^\dagger + V_{l'l}^\dagger I(\nu, \vec{d})) d\nu d\Omega \\ - n_l \sum_{l' \neq l} C_{ll'} - n_l \sum_{l' \neq l} \oint \int \frac{1}{h\nu} (U_{ll'}^\dagger + V_{ll'}^\dagger I(\nu, \vec{d})) d\nu d\Omega = 0. \end{aligned} \quad (2.30)$$

U and V are marked with daggers so as to later incorporate the necessary PRD effects. [Rybicki & Hummer \(1992\)](#) defined a new operator  $\Psi$  such that

$$\Psi_{\nu, \vec{d}}[\eta] = \Lambda_{\nu, \vec{d}}[(\chi^\dagger)^{-1} \eta] \quad (2.31)$$

These two operators,  $\Lambda$  and  $\Psi$  are equivalent for a converged solution as  $\chi^\dagger = \chi$ , but the use of  $\Psi$  is necessary to obtain the form of statistical equilibrium equations preconditioned to be linear in the atomic populations that is core to this method.

## 2.4 Introduction to Radiative Transfer

The operator splitting technique is again applied here. We then have

$$I(\nu, \vec{d}) = \Psi_{\nu, \vec{d}}^*[\eta(\nu, \vec{d})] + (\Psi_{\nu, \vec{d}} - \Psi_{\nu, \vec{d}}^*)[\eta^\dagger(\nu, \vec{d})], \quad (2.32)$$

and then considering the effects on one atom, under the assumption that background emissivity and opacity do not change during an iteration

$$I(\nu, \vec{d}) = I^\dagger(\nu, \vec{d}) - \sum_j \sum_{i < j} \Psi_{\nu, \vec{d}}^* [n_j^\dagger u_{ji}^\dagger] + \sum_j \sum_{i < j} \Psi_{\nu, \vec{d}}^* [n_j u_{ji}^\dagger], \quad (2.33)$$

where  $i$  and  $j$  refer to the levels present in the atomic model. The first two terms of this expression are often termed  $I^{\text{eff}}$  and in the case of a diagonal  $\Psi^*$  operator this represents the non-local contribution to the radiation field from the current atom, and the contribution from all other species.

We can write the preconditioned statistical equilibrium system (2.30) as

$$\Gamma \vec{n} = \vec{0}, \quad (2.34)$$

where  $\Gamma$  is a matrix consisting of the sum of  $\Gamma^R$  due to the radiative contributions, and  $\Gamma^C$  from the collisional contributions.  $\vec{n}$  is the vector of the updated level populations for the species at this point in the atmosphere. We can then write

$$\Gamma_{ll'}^R = \oint \int \frac{1}{h\nu} \left( u_{l'l}^\dagger + v_{l'l}^\dagger I_{\nu, \vec{d}}^{\text{eff}} - \left( \sum_{m \neq l} \chi_{lm}^\dagger \right) \Psi_{\nu, \vec{d}}^* \left[ \sum_p u_{l'p}^\dagger \right] \right) d\nu d\Omega \quad (2.35)$$

for  $l \neq l'$ . As the collisional rates can be treated directly we have simply

$$\Gamma_{ll'}^C = C_{l'l}, \quad (2.36)$$

for  $l \neq l'$ . The problem is now represented by a system of equations linear in the level populations. Due to the necessity of total number conservation the sum of the entries in each column of  $\Gamma$  must be 0, which allows us to compute the diagonal entries of the complete  $\Gamma$  matrix as

$$\Gamma_{ll} = - \sum_{m \neq l} \Gamma_{ml}. \quad (2.37)$$

An additional constraint must be applied to (2.34) to determine the new populations, and avoid the trivial solution of  $\vec{n} = \vec{0}$ , which is otherwise always a valid solution. This constraint is typically expressed through local population conservation, which amounts

to replacing one of the rows of  $\Gamma$  with ones, and the associated entry in the right-hand-side zero-vector with the local species number density  $n_{\text{total}}$ . There is then a  $\Gamma$  matrix at each point in the atmosphere for each atomic species being considered in detail; each of these is square with dimension equal to the number of levels in the atomic model for this species. This system can then be solved for  $\vec{n}$ , independently at each location in the atmosphere, as the current estimate of non-local coupling through the radiation field is already present in the terms that compose  $\Gamma$ .

Iterating the populations through (2.34) with interleaved formal solutions gives us a reliable and rapidly converging method for solving the multilevel NLTE problem with multiple atoms and overlapping transitions as used in the *Lightweaver* framework. A more detailed description of the numerical implementation is provided in the *Lightweaver* paper (Osborne & Milić 2021).

#### 2.4.11 Time-Dependent Population Updates

As discussed previously, an iterative update to the populations in the statistical equilibrium case can be phrased as  $\Gamma\vec{n} = \vec{0}$ . A variant of this method can also be applied to the time-dependent form of the kinetic equilibrium equations (2.11). We will not directly treat the advection term of equation (2.11), as this requires consideration of hydrodynamic effects and the discretisation schemes used therein (due to the large gradients of these populations that occur in the solar atmosphere)<sup>6</sup>. We can, however, discretise  $\partial n/\partial t = \Gamma\vec{n}$  as

$$\frac{\vec{n}^{t+1} - \vec{n}^t}{\Delta t} = \theta\Gamma^{t+1}\vec{n}^{t+1} + (1 - \theta)\Gamma^t\vec{n}^t, \quad (2.38)$$

where  $\theta$  indicates the degree of implicitness, the  $t$  and  $t + 1$  indices indicate the start and end of the timestep being integrated respectively, and  $\Delta t$  the duration of the timestep.  $\vec{n}^{t+1}$  can then be found by rewriting (2.38) as

$$(\mathbb{I} - \theta\Delta t\Gamma^{t+1})\vec{n}^{t+1} = (1 - \theta)\Delta t\Gamma^t\vec{n}^t + \vec{n}^t, \quad (2.39)$$

with  $\mathbb{I}$  the identity matrix. This system is then iterated until  $\vec{n}^{t+1}$  converges, for a new evaluation of  $\Gamma^{t+1}$  at each iterate.

---

<sup>6</sup>An overview of the methods used to numerically solve the conservation laws of hydrodynamics is provided in Sec. 2.5.

### 2.4.12 Partial Frequency Redistribution

Most solar spectral lines form in regions where complete frequency redistribution (CRD) holds. That is to say that the plasma is sufficiently collisional that elastic collisions redistribute electrons across all sub-states of an energy level prior to emission. In this case, the emission frequency of a photon is not correlated with the frequency of the photon absorbed to excite the atom into this state i.e. photons are completely redistributed in frequency and the line emission and absorption profiles are equal. In lower density regions with strong, typically resonance<sup>7</sup>, lines where radiative effects dominate over collisional effects, a portion of the population of each level will emit photons with a frequency correlated with the previous absorption frequency. The remainder of the level population is said to be *natural*, with an emission frequency independent of the manner in which the level was populated. (Hubený & Mihalas 2014). The line's emission and absorption profiles then differ and this coherent scattering must be treated explicitly. This imposes substantial computational effort, but we will briefly describe the key points of the process, and how it fits into the MALI framework following Uitenbroek (2001) and Hubený & Mihalas (2014).

We can define the emission profile coefficient  $\rho_{ij} = \psi_{ij}/\phi_{ij}$ , at which point all  $U$  and  $V$  terms can be rewritten in terms of  $\rho_{ij}$  and  $\phi$ , and  $\dagger$  on these terms refers to the value of  $\rho_{ij}$  evaluated at the previous iterate (analogous to  $n^\dagger$ ).

From this definition of  $\rho_{ij}$ , under the assumptions of a line with an infinitely sharp lower level and broadened upper level, and the validity of PRD in the atomic frame being approximated by PRD in the observer's frame (Uitenbroek 2001), following Hubený & Mihalas (2014) we have

$$\rho_{ij}(\nu, \vec{d}) = 1 + \gamma \frac{\sum_{l < j} n_j B_{lj}}{n_j P_j} \oint \frac{1}{4\pi} \int I(\nu', \vec{d}') \cdot \left[ \frac{R_{lji}^{II}(\nu', \vec{d}'; \nu, \vec{d})}{\phi_{ij}(\nu, \vec{d})} - \phi_{ij}(\nu', \vec{d}) \right] d\nu' d\Omega', \quad (2.40)$$

where  $R^{II}$  is the generalised redistribution function for transitions of this kind (Hubený 1982),  $\gamma$  is the coherency fraction, and  $P_j$  is the total depopulation rate of level  $j$ . The  $lji$  subscript on  $R^{II}$  describes the scattering process, indicating that electrons can start from any level  $l$  in the range  $[i, j)$ . These scattering processes are then summed. The cross-redistribution, or Raman scattering, processes for which  $l \neq i$  are typically less

<sup>7</sup>A transition whose lower state is the ground state of the atom is known as a resonance transition.

## 2.4 Introduction to Radiative Transfer

important than resonance scattering ( $l = i$ ). Only considering the latter of these simplifies the evaluation of  $\rho_{ij}$ . The redistribution function describes the probability that a photon with frequency  $\nu'$  and direction  $\vec{d}'$  is re-emitted with frequency  $\nu$  and direction  $\vec{d}$ .

The coherency fraction  $\gamma$  describes the probability of a photon being emitted from its current sublevel of the energy level  $j$ , prior to an elastic collision that would redistribute it across the sublevels of  $j$ . This is computed as

$$\gamma = \frac{P_j}{P_j + Q_j}, \quad (2.41)$$

where  $Q_j$  is the total rate of elastic collisions affecting level  $j$ .

Ignoring bulk plasma motions that create anisotropy in the radiation field, the integrals over angle and frequency can be split, and render the calculation of this angle-averaged form of  $\rho_{ij}$  much simpler and less computationally demanding. Also defining  $g_{II}(\nu, \nu') = R^{II}(\nu, \nu')/\phi_{ij}(\nu')$  such that the fast approximation of [Gouttebroze \(1986\)](#) and [Uitenbroek \(1989\)](#) can be employed, and ignoring cross redistribution terms we have

$$\rho_{ij}(\nu) = 1 + \gamma \frac{n_i B_{ij}}{n_j P_j} \left( \int g_{II}(\nu, \nu') J(\nu') d\nu' - \bar{J}_{ij} \right), \quad (2.42)$$

where

$$\bar{J}_{ij} = \frac{1}{4\pi} \oint \int I(\nu, \vec{d}) \phi(\nu, \vec{d}) d\nu d\Omega = \frac{R_{ij}}{B_{ij}} \quad (2.43)$$

is the frequency-integrated mean intensity across the transition. In the case of plasma flows we adopt the approximate hybrid treatment of [Leenaarts et al. \(2012b\)](#), using  $J$  in the plasma's rest frame, to compute  $\rho_{ij}$  in this same frame. Its value is then interpolated to find the value of  $\rho_{ij}$  at the Doppler shifted value in the observer's frame. The additional computational cost of this hybrid method over the angle-averaged method is very low, and in most cases the results are comparable to the far more costly complete angle-dependent treatment ([Leenaarts et al. 2012b](#); [Kerr et al. 2019a](#)).

We adopt the iterative method of [Uitenbroek \(2001\)](#) to evaluate  $\rho_{ij}$  by interleaving formal solutions at wavelengths of PRD lines (and wavelengths that are Doppler shifted into the region of a PRD line in the case of hybrid PRD) between updates of  $\rho_{ij}$  from the current value of the radiation field. The atomic populations are held constant during this process, and a handful of iterations of computing  $\rho_{ij}$  and updating the radiation field are performed between each MALI iteration.

### 2.4.13 Charge Conservation

We have thus far assumed that the electron density is known *a priori* as part of the atmospheric model. Unfortunately, it is also insufficient to assume that the electron density follows the LTE ionisation state of the plasma, as many species (in particular hydrogen) with important NLTE spectral lines will be far from their LTE ionisation state (e.g. [Heinzel 1995](#); [Paletou 1995](#); [Bjrgeren et al. 2019](#)). Synthesis of lines with an incorrect electron density will often converge to different final populations, and produce different spectral line shapes. A secondary iteration process is needed to determine the electron density in self-consistent way within the MALI framework. This is achieved through a Newton-Raphson iteration first proposed by [Heinzel \(1995\)](#) and [Paletou \(1995\)](#). The method presented here differs slightly from that presented by these authors as we choose to include the effects of all species considered in NLTE in the calculation of charge conservation.

The statistical equilibrium equations of level  $i$  of species  $s$  can be written as a function of the species' level populations and the electron density

$$F_{s,i}(\vec{n}_s, n_e) = \sum_{j \neq i} n_j P_{ji}(\vec{n}_s, n_e) - n_i \sum_{j \neq i} P_{ij}(\vec{n}_s, n_e) = 0. \quad (2.44)$$

Under our previous assumption of fixed electron density this expression is linear in unknown populations and reduces to the preconditioned system of the MALI method (2.34). We first obtain the solution to this system and denote this intermediate result  $\vec{n}_s^{\sim}$ .

A Newton-Raphson iteration is applied to this system to compute the correction to the level populations and electron density that will drive the  $F_{s,i}(\vec{n}_s, n_e)$  towards 0. This is achieved by using the Jacobian of  $F$ . For an arbitrary function  $f(\vec{x})$ , to compute the correction  $\delta\vec{x}$  to an initial parameter  $\vec{x}_0$  such that  $f(\vec{x}_0 + \delta\vec{x}) = 0$  it is sufficient to solve  $-J\delta\vec{x} = f(\vec{x}_0)$  for  $\delta\vec{x}$ , where  $J$  is the Jacobian of  $f$  evaluated at  $\vec{x}_0$ . Applying this procedure to the preconditioned statistical equilibrium equations gives

$$F_{s,i}(\vec{n}_s^{\sim}, n_e^{\dagger}) = - \sum_j \left( \frac{\partial F_{s,i}(\vec{n}_s, n_e)}{\partial n_j} \Big|_{(\vec{n}_s^{\sim}, n_e^{\dagger})} \delta n_j \right) - \frac{\partial F_{s,i}(\vec{n}_s, n_e)}{\partial n_e} \Big|_{(\vec{n}_s^{\sim}, n_e^{\dagger})} \delta n_e, \quad (2.45)$$

where  $\delta n_j$  and  $\delta n_e$  are the necessary corrections to obtain self-consistent level populations and electron density.

This system applies *simultaneously* to all species to be considered in the determination of the self-consistent electron density, and contains  $\sum_s N_{\text{level},s} + 1$  equations, where  $N_{\text{level},s}$  is the number of levels in the model of species  $s$ . Similarly to the initial population update

## 2.4 Introduction to Radiative Transfer

through the MALI method (2.34), constraints are needed to ensure that all parameters are correctly conserved. Here we require constraints on the population of each species, and on the charge neutrality of the system. The former of these is expressed as

$$\sum_j \delta n_j = n_{\text{total}} - \sum_j \tilde{n}_j, \quad (2.46)$$

for each species, while the latter is described by

$$\delta n_e - \sum_s \sum_j \text{ion}_s(j) \delta n_{s,j} = n_{e,\text{bg}} + \sum_s \sum_j \text{ion}_s(j) - n_e^\dagger, \quad (2.47)$$

where  $\text{ion}_s(j)$  is the ionisation state of the  $j$ -th level of species  $s$ , and  $n_{e,\text{bg}}$  represents the electron density due to background species not considered in detail here. The left-hand side of the population conservation equation replaces one of the rows of the Jacobian for each species, analogously to the population conservation equation used in (2.34). The charge conservation equation is the “extra” equation in this system, and the left-hand side of the expression here forms one row of the Jacobian.

The final Jacobian matrix is therefore block diagonal, with one row that couples these blocks to each other. The left-hand side of the system (2.45) is given by the right-hand sides of the previous constraint equations, and  $F_{s,i}(\tilde{n}_s, n_e^\dagger) = \Gamma_s \tilde{n}$  for the remaining entries for each species  $s$ . This system can now be solved in the same fashion as (2.34), and the final populations can be computed from  $n_j = \tilde{n}_j + \delta n_j$  and  $n_e = n_e^\dagger + \delta n_e$ . The exact form of these derivatives used in the Jacobian can be computed analytically for all rates, and these terms are presented in Osborne & Milić (2021). The derivatives with respect to the level populations are simply the associated entries of  $\Gamma$  for each species, whilst those associated with the electron density depend on the collisional and bound-free radiative rates, but we do not find the exact form to be particularly insightful.

The time-dependent case can be solved very similarly to the previous statistical equilibrium case. Following Kašparová et al. (2003), and starting from the time-dependent population update equation (2.39), we can define

$$G_{s,i}(\tilde{n}_s^{t+1}, n_e) = n_{s,i}^{t+1} - \theta \Delta t F_{s,i}(\tilde{n}_s^{t+1}, n_e) - (1 - \theta) \Delta t \Gamma_s^t \tilde{n}_s^t - n_{s,i}^t = 0, \quad (2.48)$$

and apply this secondary Newton-Raphson iteration process to  $G$ , using the same constraint equations as before.

It is important to stress that only one Newton-Raphson iteration is needed following each

standard preconditioned population update, as the original MALI system was already preconditioned and linear in the populations. The iterative process proceeds by following each MALI population update with a secondary Newton-Raphson iteration evaluated using the updated  $\Gamma$ , and then returning to the evaluation of  $\Gamma$ , or  $\rho_{ij}$  in the presence of transitions treated with PRD.

#### 2.4.14 Determining Numerical Convergence

The iterative procedures presented in the previous sections yield refined estimates of the level populations throughout the model atmosphere for each additional iteration. This raises the question of when to stop iterating, and by what metric to measure convergence. Convergence describes the difference between these values for two subsequent iterations, and how far they are from the true values for the model. As we do not generally have knowledge of the latter, the former is typically used to determine when to cease iterations. The most important quantities to use when estimating convergence are the populations  $n$ , and radiation field  $J$ , due to their appearance in the source function.

In NLTE problems, where spectral lines will often form in relatively compact regions of the atmosphere, the change may only be large in this region, so the  $L^1$  or  $L^2$  norms of the difference between successive iterates may not be very informative. Additionally, the value of both  $J$  and  $n$  will vary hugely throughout the atmosphere, thus the absolute change over an iteration may not be particularly meaningful. Instead, it is common to use the  $L^\infty$  norm<sup>8</sup> of the relative change of these parameters. For a population  $n_i$ , following [Auer et al. \(1994\)](#), we denote the relative change  $R_c(\text{itr}, i)$  for level  $i$  and iteration  $\text{itr}$ . It can be expressed as

$$R_c(\text{itr}, i) = \max_k \left| \frac{n_i^{\text{itr}} - n_i^{\text{itr}-1}}{n_i^{\text{itr}}} \right|, \quad (2.49)$$

where  $k$  represents the location in the atmosphere, and  $n_i^{\text{itr}}$  represents the population of level  $i$  after iteration  $\text{itr}$ . By iterating until  $R_c(\text{itr}, i)$  is low for all transitions, we ensure that the largest update to a term is small relative to its current value. When applied to  $J$ , the  $L^\infty$  norm of the relative change is typically computed across frequency and atmospheric depth.

[Auer et al. \(1994\)](#) and [Fabiani Bendicho et al. \(1997\)](#) provide a framework for estimating both the error from the fully converged solution, and the truncation error due to the discretisation onto the numerical grid, but in practice this is rarely employed outside

<sup>8</sup>The  $L^\infty$  norm of a vector is given by the maximum absolute value present in its components.

## 2.5 Introduction to Hydrodynamics and Conservation Laws

of multi-grid strategies (e.g. Fabiani Bendicho et al. 1997; Léger et al. 2007). Instead, we apply the more common technique of iterating until the maximum value of  $R_c$  for the populations drops below a certain threshold (typically  $\sim 10^{-3}$ )<sup>9</sup>. This approach has been applied with success to many modelling problems within the MULTI, RADYN, and RH (Uitenbroek 2001) codes (amongst others), but we note, following Auer et al. (1994) and Fabiani Bendicho et al. (1997), that a small value of  $R_c$  does not guarantee convergence, whilst methods such as the multi-grid one they present can do so. Unfortunately, these methods are significantly more complex than the MALI method presented here and appear to have difficulty in more realistic multi-dimensional models (J. Štěpán, *private communication*).

## 2.5 Introduction to Hydrodynamics and Conservation Laws

*The majority of the basic theory in this section follows the two texts by LeVeque (1997, 2002).*

In the previous sections we have explained the basis of the radiative transfer methods used throughout this work. The difficulty in radiative transfer primarily lies in the nuance of implementing the various integration terms. To provide a clear understanding of RHD we also need a numerical description of hydrodynamics, and an explanation of solving the coupled systems of partial differential equations (PDE) that represent the other major facet of RHD.

The scalar radiative transfer equation, solved via the formal solver, is a good example of an ordinary differential equation. It is relatively easy to solve via a variety of methods, typically striving for a balance of speed, reliability, and accuracy. Unfortunately, PDEs are difficult to solve numerically in a general way.

A generic second-order PDE of a quantity  $q$  depending on two independent variables  $x$  and  $y$  can be written as

$$a q_{xx} + b q_{xy} + c q_{yy} + d q_x + e q_y + f q = g, \quad (2.50)$$

wherein the subscripts represent partial derivatives with respect to these variables. For convenience we will retain this convention through the current section. The sign of the discriminant ( $\Delta = b^2 - 4ac$ ) of this equation determines the class of the problem:

---

<sup>9</sup>It is wise to also track  $R_c$  for  $J$ , and although this is less commonly used as a convergence criterion, we often choose to do so as a safety factor. The  $R_c$  of one of  $n$  or  $J$  being significantly larger than the other likely indicates that the spatial, frequency, or angular quadrature is ill-suited to the problem at hand.

## 2.5 Introduction to Hydrodynamics and Conservation Laws

- $\Delta < 0$ : Elliptic problem (e.g. Poisson equation).
- $\Delta = 0$ : Parabolic problem (e.g. Heat equation).
- $\Delta > 0$ : Hyperbolic problem (e.g. Advection equation).

Here, we will focus primarily on hyperbolic problems, with a brief discussion of parabolic terms. Hyperbolic problems take the form

$$q_t + f(q)_x = 0, \quad (2.51)$$

where  $f$  is a function describing the flux of  $q$  at each location in the domain,  $t$  indicates a temporal coordinate, and  $x$  a spatial coordinate. In the following discussions, there is an inherent assumption that the flux function is local, and acts only on the local state variables  $q$ . This is not necessarily the case with radiative terms, but these can be incorporated into such a scheme by determining their local effects on the plasma energy.

Equations of the form (2.51) are known as conservation laws, and the quantities  $q$  are often termed “conserved quantities”, implying that  $\int_{-\infty}^{\infty} q \, dx$  is constant in time. This does not preclude the addition of sources and sinks of this quantity, but simply requires that the total of a conserved quantity not change *without* being acted on in such a way. The simplest equation of this form is advection, which arises from conservation of mass in a moving fluid and is written for mass density  $\rho$  and fluid velocity  $v$  as

$$\rho_t + (\rho v)_x = 0. \quad (2.52)$$

The advection equation can be augmented with two further equations to form the Euler equation set. These three equations are the conservation of mass, momentum, and energy. Together they describe the evolution of ideal fluids. The complete set is written

$$\begin{aligned} \rho_t + (\rho v)_x &= 0, \\ (\rho v)_t + (\rho v^2 + p)_x &= 0, \\ E_t + (v(E + p))_x &= 0, \end{aligned} \quad (2.53)$$

where  $E$  is the total energy and  $p$  is the gas pressure. This system is already recognisable as the simplified roots of the RHD equations. As our three conserved quantities are the mass density, momentum density, and energy, the pressure must be expressed as a combination of these so as to be able to write this system in the form of (2.51). The simplest way to

## 2.5 Introduction to Hydrodynamics and Conservation Laws

achieve this is to use the equation of state for an ideal gas

$$e = \frac{p}{\gamma - 1}, \quad (2.54)$$

with  $e$  the internal energy, and  $\gamma$  is the ratio of gas' specific heat at constant pressure and constant volume, which is 5/3 for monatomic gases, typically assumed in the case of solar plasma. This is related to the total energy  $E$  by

$$E = e + \frac{1}{2}\rho v^2. \quad (2.55)$$

The Euler equations (2.53) describe the evolution of an ideal fluid without any energy losses. In practice we often need to model some loss terms. Ignoring radiative effects for now, the most important of these is heat conduction<sup>10</sup>, which is typically described by parabolic equations. This adds a second spatial derivative term to the right-hand side of the energy conservation equation. It is also common to need source and sink terms when modelling real world problems; these can account for fluids entering and leaving a volume, the non-local emission and absorption of energy, or simply the effects of gravity. These terms are also added to the right-hand side of the equations of (2.53), and together with the effects of viscosity, these describe the Navier-Stokes equations.

### 2.5.1 Numerical Approaches

The typical first choice for numerically solving differential equations is to apply a finite difference method. In this class of method, the problem is discretised, similarly to the approach taken in radiative transfer, and the values associated with each grid point represent the local pointwise value. The local gradient of the conserved quantities can then be estimated from pointwise difference in the quantity between adjacent cells. This method can be applied to discrete problems in both space and time. Applying a basic one-sided method to the advection equation (2.52) gives

$$\frac{q_i^{t+1} - q_i^t}{\Delta t} + v \left( \frac{q_{i+1}^t - q_i^t}{\Delta x} \right) = 0, \quad (2.56)$$

where the subscript refers to the location of the conserved quantity, the superscript the discrete timestep, with  $\Delta x$  and  $\Delta t$  being the local grid spacing and timestep duration

---

<sup>10</sup>In the plasmas considered here there is additional nuance to heat conduction, which we will return to in Sec. 2.6.

respectively.

It is clear that without loss of generality we could have also chosen to spatially difference our problem in the other direction (i.e.  $q_i^t - q_{i-1}^t$ ). From the infinitesimal definition of the derivative, these two formulations are equivalent. This is not the case in discretised problems. It is better to locally use the formulation such that information from points “up-wind” in terms of the fluid velocity are used to update the points downwind of themselves. In this sense the information used to update points is following the fluid flow. Equation (2.56) can be rewritten to explicitly determine the approximate value of the quantity at the next timestep,

$$q_i^{t+1} = q_i^t - \frac{v\Delta t}{\Delta x} (q_{i+1}^t - q_i^t). \quad (2.57)$$

This simple first order accurate approach can be applied to any conservation law, and higher-order accurate methods can be derived from using finite difference methods over larger stencils, or deriving similar approaches from combinations of the local finite difference approximations using Taylor series.

Many other spatial and temporal discretisations can be devised for conservation laws, and it is important to choose a discretisation that introduces minimal error. In general, we term discretisations where  $q^{t+1}$  depends only on  $q^t$  as *explicit*, and those also depending on  $q^{t+1}$  as *implicit*, necessitating the solution of a system of typically non-linear equations.

Whilst the finite difference method provides an intuitive formulation for discretising these equations, it is difficult (but entirely possible) to ensure conservation of the quantities that we desire be conserved with only pointwise values and no formal description of the variation between these points. Instead, the finite volume description is often preferable and consists of treating  $q_i^t$  as the average value over grid cell  $i$ . Conservation can then be ensured by evolving this value based on the fluxes in and out of the cell. This implies

$$q_i^t \approx \frac{1}{\Delta x} \int_{x_i}^{x_{i+1}} q^t(x) dx. \quad (2.58)$$

Rewriting the conservation law (2.51) in an integral form then gives

$$\int_{x_i}^{x_{i+1}} q^{t+1}(x) dx = \int_{x_i}^{x_{i+1}} q^t(x) dx + \int_{t_t}^{t_{t+1}} f(q(x_i, t)) dt - \int_{t_t}^{t_{t+1}} f(q(x_{i+1}, t)) dt, \quad (2.59)$$

which can then be written as the flux-differencing form

$$q_i^{t+1} = q_i^t - \frac{\Delta t}{\Delta x} (F_{i+1}^t - F_i^t), \quad (2.60)$$

## 2.5 Introduction to Hydrodynamics and Conservation Laws

where  $F_i$  represents the flux due between cells  $i$  and  $i - 1$ , and  $F_{i+1}$  the flux between cells  $i$  and  $i + 1$ . In the case of an explicit method with a correctly chosen timestep for the grid, where information cannot move further than one cell in a timestep, each of these flux functions depends only on cell  $i$  and one of its neighbours. An important strength of this method for solving conservation law problems is that if  $F_i$  and  $F_{i+1}$  are respectively the left and right edge fluxes for cell  $i$ , then  $F_{i+1}$  will be the left edge flux for cell  $i + 1$ , and  $F_i$  will be the right edge flux for cell  $i - 1$ . Thus the numerical integration of  $q^{t+1}$  is conserved with respect to  $q^t$ , as all numerical fluxes, other than the left- and right-most, will cancel due to the formulation of (2.60). These left- and right-most fluxes will need to consider the boundary conditions of the finite simulation volume to ensure the correctness of the conservation law. Equivalent formulations can be found for finite-difference approaches, but are slightly harder to arrive at.

### 2.5.2 Riemann Problems

The finite volume method we have described therefore depends on finding expressions for the numerical flux between two adjacent cells. This is often framed as a Riemann problem at the cell interface. A Riemann problem consists of solving a conservation law with an initial parameter distribution consisting of two constant states meeting at a discontinuity.

This problem can be analytically solved by considering the propagation of waves from this interface. From analysis of the Jacobian of the flux function of the Euler equations, the structure of this solution can be revealed. There are three waves, associated with the eigenvalues of this Jacobian:  $v$  and  $v \pm c_s$ , where  $c_s$  is the sound speed of the fluid. These waves can be non-linear and do not necessarily propagate at the characteristic velocity given by their associated eigenvalue. The expected solution is a contact discontinuity propagating with the fluid velocity  $v$ , and two non-linear waves, a shock wave and a rarefaction wave. That states on either side of these waves will need to be computed using an iterative process after the application of the Rankine-Hugoniot jump conditions.

As the complexity of the system and the equation of state increases, it becomes harder (or impossible) to compute this solution in an efficient manner. Fortunately, it is rarely necessary to compute the exact solution to the Riemann problem as many efficient approximate solvers exist, but the structure originating from this simple case can be used to interpret and validate other numerical methods. This solution will need to be computed at every interface to determine the fluxes between the cells, but the question remains of

how to define the states left and right of the interface in the definition of the Riemann problem.

### 2.5.3 Godunov's Method and Higher Order Reconstructions

The method of [Godunov \(1959\)](#) consists of assuming that the data  $q$  is piecewise constant in each cell of our simulation domain. At this point the values on the left- and right-hand side of each interface are known and the flux through this interface can be computed. The Riemann problem at each interface can be treated independently under the assumption that the fastest wave from one interface not carry information to the next. This is a fundamental requirement of stability and will be discussed in more detail in Sec. 2.5.5.

This method provides a basic framework for solving conservation laws. It is limited by the assumption that the data is piecewise constant, but nevertheless paved the way for some of the most accurate numerical methods for conservation laws. [van Leer \(1979\)](#) provided one of the first higher order extensions to Godunov's method. It is tempting to attempt to estimate the value of  $q$  at the cell interfaces with higher accuracy by using some form of reconstruction (a method closely related to interpolation), but care must be taken with this approach. The Monotonic Upstream-centred Scheme for Conservation Laws (MUSCL) method of [van Leer \(1979\)](#) uses a piecewise linear approach to reconstruction in each grid cell, but limits the gradient of the reconstruction to prevent the addition of under- or over-shoots to the data. This is achieved through use of a slope-limiter such that a monotonic series of cell averages be preserved by ensuring that the reconstructed slope not take values beyond the average of the adjacent cells. If the current cell is an extremum then the slope is set to 0. This method tracks the data to second-order in regions of smooth variation, but degenerates to the first-order Godunov method at discontinuities. The reconstruction technique used in RADYN is based on this method.

There are further high-order extensions to the concept of reconstruction including the parabolic method of [Colella & Woodward \(1984\)](#) providing third-order accuracy in smooth regions, and the general weighted essentially non-oscillatory (WENO) methods. WENO methods are a cornerstone of reconstruction<sup>11</sup> in modern finite volume and finite difference codes, and as such we will describe them briefly here.

---

<sup>11</sup>The properties that make WENO methods a good choice for reconstruction also render them applicable to interpolation. Throughout *Lightweaver* and the other numerical tools presented in this thesis we make use of a fourth-order WENO interpolation method described by [Janett et al. \(2019\)](#) for its accuracy in smoothly varying regions and reliable behaviour around sharp variations. For example, in *Lightweaver* it is used for the interpolation of photoionisation cross-sections onto the final wavelength grid.

## 2.5 Introduction to Hydrodynamics and Conservation Laws

WENO methods were first proposed by Liu et al. (1994), and formalised for arbitrary order by Jiang & Shu (1996). These methods form a convex combination of polynomials over overlapping regions. For example, in the case of the commonly used fifth-order WENO method of Jiang & Shu (1996), three parabolae are constructed over five adjacent points, i.e. each using three contiguous points. We can equivalently define a fourth-order polynomial over the five points, which can also be written as a linear combination of our interpolating parabolae at each point in this region. These linear weights are further weighted by a term known as the *smoothness indicator*, which estimates the local smoothness, such that the weight of each term is equal to its expected linear weight in smooth regions, and heavily biased towards a parabola in a smooth region if other regions present discontinuities. WENO methods are highly performant thanks to the lack of conditional branches in this core procedure.

In the case of the finite volume method these functions can be derived to reconstruct the interface values from the cell average values, and for a fixed uniform grid, the integration weights for each interface can be computed analytically. For non-uniform grids, the weights can be precalculated if the grid remains fixed, or computed on the fly if necessary. There are also a number of approximate methods for handling non-uniform grids at low computational cost such as WENO-NM (Huang et al. 2018).

### 2.5.4 Numerical Fluxes

Whilst there are many approximate Riemann solvers that can accurately and efficiently solve the Riemann problems arising from the Euler equations including the addition of source terms, it does not appear feasible to express the evolution of the entire RHD system in this way. Instead, in explicit methods we can use numerical approximations to the flux based on the reconstructed values at the interfaces. For implicit methods we can instead use a Newton iterative scheme to minimise the residual of the discretised conservation law across the cell interfaces with the fluxes typically computed from the upwind value of the reconstructed parameter.

With high-order reconstructions, simple expressions for the fluxes can be used and give accurate results. An obvious first choice would be the average flux from the reconstructed states left and right of the interface. Unfortunately, this leads to numerical instabilities, and some artificial damping (numerical viscosity) is needed. This leads to a flux known as

the Local Lax-Friedrichs, or Rusanov Flux ([Rusanov 1962](#))

$$F_{\text{LLF}} = \frac{1}{2}(f(q_{i+1}^{\text{L}}) + f(q_i^{\text{R}})) - \frac{1}{2}\alpha(q_i^{\text{R}} - q_{i+1}^{\text{L}}). \quad (2.61)$$

Here  $q_i^{\text{L}}$  and  $q_i^{\text{R}}$  represent the left- and right-hand reconstructed states of the  $i$ -th Riemann problem,  $f$  represents the flux function of the conservation law, and  $\alpha$  is the local maximum wave propagation speed for this system. Whilst  $\alpha$  can often be formally derived from the Jacobian of  $f$ , it can often be replaced with either the maximum absolute value of the sum of the sound speed on each side of the interface and the local fluid velocity, or the average of these on both sides of the interface. General symmetric fluxes like this are simple and efficient, but tend to be diffusive, smearing features across many grid cells, especially if the first-order Godunov or van Leer-style reconstruction methods are used. This can be effectively combatted by the use of high-order reconstruction schemes and adaptive mesh refinement techniques, providing simple and robust solutions thanks to the reconstruction scheme.

### 2.5.5 Time Integration, Stability, and Splitting Schemes

Explicit schemes are only stable (i.e. do not diverge or introduce spurious oscillations) if the Courant-Friedrichs-Lewy (CFL) condition is met. The CFL condition states that the numerical domain of dependence of the equation (i.e. the terms used in the computation of a value in the next timestep) must encompass the analytic domain of dependence to ensure that all necessary information is taken into account. This is a necessary, but not sufficient condition, and the exact requirements of each scheme can often be derived analytically, although it is common to apply an additional safety margin to the maximum permitted value of the CFL condition.

For explicit methods the CFL condition typically sets the maximum timestep that can be used based on the current simulation conditions. In a hyperbolic system the CFL condition takes the form

$$C = \frac{v\Delta t}{\Delta x}, \quad (2.62)$$

and will be constrained to a maximum value ( $\leq 1$ ) for the stability of a low order explicit method. Most implicit methods do not place an upper limit on the CFL condition for stability (as all points are coupled), but typically need to remain  $\sim 1$  to avoid losing fine detail in the solution. This is discussed by [Viallet et al. \(2011\)](#), who comment that  $\text{CFL} \sim 1$  serves

## 2.5 Introduction to Hydrodynamics and Conservation Laws

as an accuracy criterion and typically represents the optimum accuracy/computational cost ratio despite the method being nominally stable for large CFL values.

It is difficult to achieve better than second order accuracy due to the temporal discretisations discussed so far, but by viewing  $F_i^t$  as the flux at time  $t$ , we can achieve higher order accuracy by applying a multi-step method to more accurately integrate these fluxes (and any associated source terms) over time. A good choice for this is a method in the family of total-variation diminishing Runge-Kutta methods (e.g. [Shu & Osher 1988](#)). These multi-step time integration methods can be combined with fractional step methods that allow us to perform *operator splitting*. That is, splitting an equation into two subproblems that can be solved independently. An example of this would be the radioactive decay of an isotope transported by advection. Splitting this into an advection and a reaction problem allows for standard methods to be used in both of these problems, but clearly their results need to be coupled to each other. A naive first approach is to solve one of the subproblems over the timestep, and then solve the other, but this method cannot be better than first order accurate in time for coupled subproblems. A commonly used approach that is second order is known as Strang splitting ([Strang 1968](#)), which consists of time-advancing the first subproblem by half the timestep, then the second by the whole timestep, before once again advancing the first subproblem by half the timestep. This method can provide second-order accuracy. Many more advanced splitting procedures have been developed, but Strang splitting remains widely used due to its ease of implementation.

[LeVeque \(1997\)](#) comments that in many situations the first-order splitting described above performs better than would likely be expected from its formal first-order accuracy. This is because the errors introduced are equivalent to solving the problem at a slightly different time, differing by up to a single timestep. Whilst this renders the method first-order accurate, the quality of the solution is still primarily controlled by the quality of the methods used to solve the subproblems, and this exceedingly simple splitting scheme can often be applied with no problems.

There are many nuances to each of the techniques needed to accurately and robustly solve conservation laws numerically, and no single *correct* method to use. Indeed, formulations based on finite volume, finite difference, finite element, and discontinuous Galerkin methods are all in use in state-of-the-art research codes. The above is intended to serve as a *somewhat opinionated* introduction to the complexities of these methods, but by no means present a conclusion on how conservation laws should be solved. It is hoped that this introduction will improve general understanding of this aspect of RHD codes.

## 2.6 Conduction

When expressed in terms of energy the heat equation in a plasma along a magnetic field line is given by

$$\frac{\partial E}{\partial t} = \frac{\partial}{\partial z} \left( \kappa_0 T^{5/2} \frac{\partial T}{\partial z} \right), \quad (2.63)$$

with spatial coordinate  $z$ , and coefficient  $\kappa_0$  varying, but typically taken to be approximately  $1 \times 10^{-6} \text{ erg cm}^{-1} \text{ s}^{-1} \text{ K}^{-3.5}$  (Spitzer & Härn 1953; Braginskii 1965) based on deviations from a fully ionised hydrogen plasma. The form of this equation poses several problems, the most significant being that in regions of sufficiently steep temperature gradients, the conductive flux approaches infinity. Clearly this is not physical and there is a maximum limit, known as the free-streaming limit, at which all electrons in the plasma are flowing at their thermal speed with this heat gradient, representing a finite limit on this term. As this free-streaming limit is approached the heat flux becomes non-local and depends on the global temperature and density structure in the loop (Battaglia et al. 2009). Campbell (1984) provides an expression for the transport coefficients used to determine the conductive flux through an ionised plasma and smoothly handles both the Spitzer-Härm, locally limited, and non-locally limited regimes. This approach can be applied in numerical simulations but more frequently (such as in RADYN) the method of Fisher et al. (1985) is applied which smoothly limits the local flux to remain less than the local free-streaming limit, helping to stabilise this equation.

Due to its parabolic nature, an explicit solution of the heat equation can be extremely costly; stability requirements provide a timestep requirement scaling with the inverse square of the grid spacing, rather than linearly as in the case of most hyperbolic equations. Any attempt to explicitly integrate the heat equation on a timestep limit set by the hydrodynamic equations will likely be met with rapid divergence of any small perturbation in the data. This renders the conduction term stiff compared to the hydrodynamical terms and it may therefore be advantageous to use an implicit method to guarantee stability. Several alternative methods have been developed, such as expressing the parabolic equation as a hyperbolic wave equation (Rempel 2016), implicit-explicit methods (which may also be used for integrating stiff source terms such as the atomic level population transition rates) (e.g. Ascher et al. 1995), or accepting the cost of an explicit discretisation in exchange for simplicity and accuracy (Bradshaw & Mason 2003; Bradshaw & Cargill 2013).

Recently, discussion has emerged around the concept of turbulence suppressed conduction (Bian et al. 2016), where turbulence restricts the motion of electrons along the loop, and it has been suggested that the standard assumption of collisionally-dominated

conduction may be a significant overestimation of true conduction rates. Simulations using the zero-dimensional enthalpy based EBTEL code (Klimchuk et al. 2008) currently suggest that this turbulent suppression alone is insufficient to maintain the high coronal temperatures and slow cooling times seen in observations, but likely represents an important component of this effect (Bian et al. 2018). As part of further investigation, work is currently under way to integrate these effects in the RADYN and HYDRAD codes.

## 2.7 Discussions

We have provided an overview of the commonly used techniques for simulating solar flares and the observable radiation produced from these events. The core components of radiative transfer, hydrodynamics, and heat conduction have been discussed in depth, along with an overview of the numerical treatments of these terms. Whilst there are other important terms in the RHD equations, such as the heating model, these represent the core of any flare simulation.

In Chaps. 4–6 we will build on the theory presented in this chapter with simulations built using our *Lightweaver* framework. The *Lightweaver* radiative transfer framework is described in detail in Chap. 4 and provides components for computing the intensity and atomic level population updates (in both statistical equilibrium and time-dependent situations) as described in Sec. 2.4. It is therefore not a complete RHD code like RADYN, but provides a more thorough treatment of the radiative transfer terms through the use of the fully preconditioned MALI method with a cubic Bézier spline formal solver. The novel modularity of its design enables greater flexibility when treating the radiative terms of RHD flare modelling and revisiting existing assumptions regarding their treatment, as different methods can easily be combined and compared.

Modern RHD simulations have significantly enhanced our understanding of flares, but remain limited in the dimensionality of their treatments and cannot currently produce plausible synthetic light-curves without the artificial superposition of many individual simulations (e.g. Kerr et al. 2020). It is likely that significant progress will be made on these limitations in the coming 1–2 decades as computing power and numerical techniques improve. In the meantime, we can undertake numerical experiments to evaluate individual components of these larger treatments (along with re-evaluating certain current assumptions) and we will present several of these using our *Lightweaver*-based exper-

## 2.7 Discussions

iments, that leverage the dynamics of RADYN RHD simulations, with the flexible and enhanced treatment of radiative transfer and level populations from *Lightweaver*.

# 3

## Optical Flare Observations and Inversions

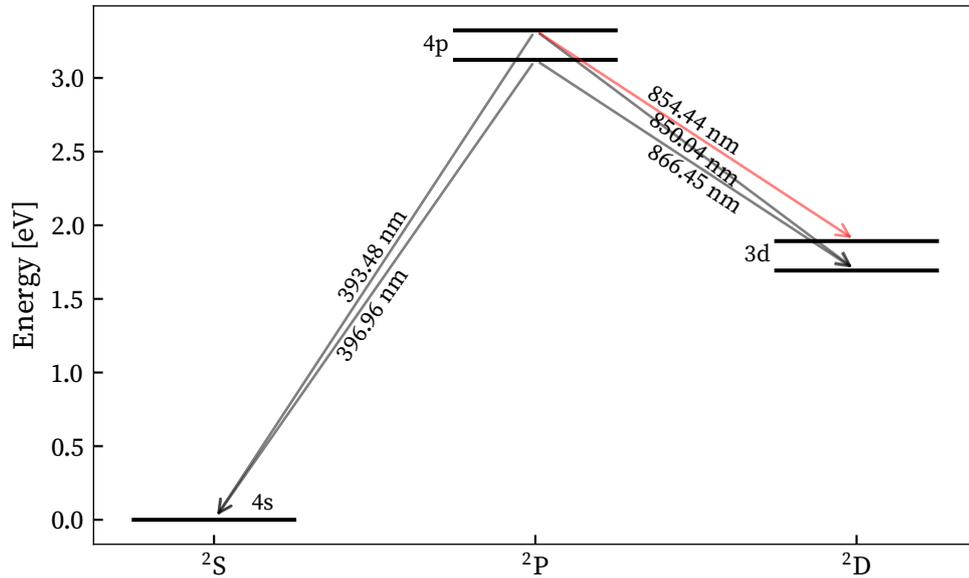
Solar flares release a vast quantity of energy across the entire electromagnetic spectrum, observed from  $\gamma$ -rays to microwaves. Almost all of this presents significant diagnostic potential, and is observed by a plethora of advanced instruments. Many of these have to be situated above the Earth's atmosphere, which shields us from the biologically harmful wavelengths at which they observe. An in-depth review of different spectral ranges, and the flaring signatures they observe is given in [Fletcher et al. \(2011\)](#).

In this chapter we will discuss the optical spectral lines that we focus on during this thesis, the basics of inversion (retrieving a solar atmosphere from spectral line observations), and finally introduce the concepts used in machine learning that will be applied in our machine learning inversion model presented in Chap. 7.

### 3.1 Important Optical Spectral Lines

Ground-based telescopes can have much larger apertures and be more complex than their space-based counterparts for an equivalent budget, allowing for much better spatial resolution. The instruments attached to them can also be repaired, replaced, and upgraded over time, allowing them to have much longer lifecycles than space missions. In exchange for this they can only observe in limited spectral bands (due to the aforementioned atmospheric absorption effects), and suffer from the effects of atmospheric turbulence. To some degree the effects of atmospheric seeing can be accounted for and corrected, but the magnitude of these effects will vary on a daily basis. Our aim is to develop predictions and techniques allowing the more thorough exploitation of the high spatially and spectrally resolved data obtained from the current and next-generation of ground-based optical telescopes to enhance our understanding of the chromosphere during solar flares.

### 3.1 Important Optical Spectral Lines



**Figure 3.1:** Grotrian (term) diagram of Ca II bound terms of interest. The Ca II 854.2 nm transition is highlighted in red. The wavelength of each bound-bound transition is labelled.

There are but a handful of spectral lines that are both accessible to these ground-based telescopes and that are sensitive to the chromosphere. Following [de la Cruz Rodríguez & van Noort \(2017\)](#) these are primarily the Ca II H & K lines, the Ca II infrared triplet,  $H\alpha$ , He I D3, and He I 1083 nm. Many other spectral lines with chromospheric diagnostic potential, such as  $Ly\alpha$ , Mg II h & k, He II 30.4 nm can only be observed from space. The decision regarding which of these to focus on depends both on the availability of observations (current and future) and the complexity of modelling. The  $H\alpha$  line at 656.3 nm has long been exploited. It has a strong line core and wide wings with a varying degree of central reversal and asymmetry (e.g. [Švestka 1966](#)). These features suggest that the line responds differently to different flares, suggesting that it has good diagnostic potential. Indeed,  $H\alpha$  has been a component of many investigations into flare dynamics and evolution (e.g. [Acton et al. 1982](#); [Heinzl et al. 1994](#); [Wang et al. 1995](#); [Kuridze et al. 2015](#); [Rubio da Costa et al. 2016](#)). Other lines in the hydrogen Balmer series can also be used, optionally in conjunction with  $H\alpha$ , such as in the work of [Capparelli et al. \(2017\)](#). The spectral lines of Ca II have long been present in the RADYN code and used for chromospheric diagnostics (e.g. [Carlsson & Stein 1992](#)), and are a mainstay of flaring chromospheric analysis (e.g. [Mein et al. 1997](#); [Cauzzi et al. 2008](#); [Kuridze et al. 2015](#); [Rubio da Costa et al. 2016](#); [Kuridze](#)

### 3.1 Important Optical Spectral Lines

et al. 2018; Vissers et al. 2021; Yadav et al. 2021), thanks in part to the polarisability of the Ca II 854.2 nm line. There have been relatively few observations of the neutral helium lines at high spatial and spectral resolution in flares (e.g. Zeng et al. 2014; Libbrecht et al. 2019), but they appear to have strong diagnostic potential, although it is likely necessary to exploit spectropolarimetric information to fully interrogate these lines (Libbrecht et al. 2019). Recently unpolarised RHD modelling efforts have started to investigate the possible dimming of these lines (Kerr et al. 2021).

In the following we shall therefore focus primarily on two of the strongest optical lines,  $H\alpha$ , and Ca II 854.2 nm, which have a long history of use within RHD modelling, and as will be discussed, can often be observed with the same instrument.  $H\alpha$  is the first of the hydrogen Balmer series, and is emitted by an electron transitioning between the  $n = 3$  and  $n = 2$  bound levels of hydrogen. Ca II 854.2 nm is part of the Ca II infrared triplet, and is highlighted in red on the Grotrian diagram in Fig. 3.1. The Ca II 854.2 nm line is the most polarisable of this triplet (as it has the largest Landé factor of the three): it is not blended with other lines, and lies relatively close to the peak of the solar spectrum, providing a high flux for easy integration. The pair of transitions on this same figure, between the  $4p^2P$  and  $4s^2S$  levels, are the Ca II H & K lines.

The  $H\alpha$  and Ca II 854.2 nm lines have been extensively used together for chromospheric analysis and carry complementary information. Simulations have found them to form at different heights within the chromosphere, for example Kuridze et al. (2015) found the  $H\alpha$  line core forming in the 1.1–1.2 Mm region of a RADYN simulation, whereas the Ca II 854.2 nm line core was formed over a much more compact region, deeper in the atmosphere, around an altitude of 0.9 Mm. Diagnostics involving both  $H\alpha$  and Ca II 854.2 nm can therefore reliably constrain chromospheric models, and are sensitive to a larger region of the solar atmosphere than is possible with just one of the two. Both of these lines have been reliably modelled without the need for additional time-consuming PRD calculations, which is not the case for the Ca II H & K lines. Whilst  $H\alpha$  can be treated in CRD, Leenaarts et al. (2012a) find that three-dimensional radiative transfer modelling is needed to correctly reproduce the expected intensity structure from RMHD models. Bjørgen et al. (2019) found that the magnitude of this effect was reduced for the brightest regions of a three-dimensional active region simulation, due to the localised high chromospheric mass density, but three-dimensional modelling remained important for regions adjacent to these bright structures. Flare models with detailed chromospheric treatment are not yet possible in three-dimensions, but with the large gradients in the radiation field that occur in flares it is likely that this effect will also be important in regions adjacent to flares (as supported by the simulations presented in Chap. 6).

### 3.1 Important Optical Spectral Lines

#### 3.1.1 The Swedish Solar Telescope

The Swedish Solar Telescope (SST, [Scharmer et al. 2003](#)) is a 1-metre refracting telescope located in the Observatorio del Roque de los Muchachos, La Palma, Spain. Behind the singlet lens, the telescope is held in a vacuum to improve image quality and a Schupmann corrector is used to compensate for chromatic aberration. The light is then sent from this corrector to an adaptive optics system that compensates for atmospheric seeing effects and on to the optical bench. This adaptive optics system allows the SST to operate at close to its diffraction limit in good seeing conditions ( $\sim 0.17''$  for H $\alpha$  and  $\sim 0.21''$  for Ca II 854.2 nm).

There are three instruments usable with the SST: TRIPPEL, CHROMIS, and CRISP. TRI-Port Polarimetric Echelle-Littrow (TRIPPEL, [Kiselman et al. 2011](#)) is a spectrograph that can simultaneously observe in three different wavelength regions (across 380–1100 nm) with a spectral resolution<sup>1</sup>  $R \approx 200\,000$ . CHROMospheric Imaging Spectrometer (CHROMIS, [Löfdahl et al. 2021](#)) and CRisp Imaging SpectroPolarimeter (CRISP, [Scharmer et al. 2008, 2019](#)) are both dual Fabry-Pérot tunable filter systems that can be operated simultaneously. CHROMIS serves the blue end of the spectrum between 380 and 500 nm, whereas CRISP operates in the 510–860 nm region. The diffraction limit is lower at the CHROMIS wavelengths, but CRISP can perform full spectropolarimetric imaging.

In this thesis we make use of one set of observations taken from the SST using the CRISP instrument, so we shall briefly describe how these are processed. This observation was performed on 2014-09-06 and captures the M1.1 flare SOL 20140906T17:09 from National Oceanic and Atmospheric Administration (NOAA) Active Region 12157 at heliocentric coordinates ( $-732''$ ,  $-302''$ ). This data is available in the F-CHROMA database<sup>2</sup> and was prepared using the CRISPRED pipeline ([de la Cruz Rodríguez et al. 2015](#)) which is responsible for image calibration, alignment, instrumental effects, and seeing restoration through the multi-object multi-frame blind deconvolution algorithm (MOMFBD, [van Noort et al. 2005](#)). MOMFBD is a post-processing phase diversity algorithm that accounts for seeing that is worse, or evolving faster, than the adaptive optics system of the telescope can cope with. It has been suggested by [Armstrong & Fletcher \(2021\)](#) that the differences between the wide- and narrowband images of flares may hinder the reconstructive abilities of the

---

<sup>1</sup>The spectral resolution  $R$  describes an instrument's ability to resolve spectral features and is defined by  $\frac{\lambda}{\Delta\lambda}$  where for a wavelength  $\lambda$ ,  $\Delta\lambda$  is the smallest difference in wavelengths that can be distinguished.

<sup>2</sup><https://star.pst.qub.ac.uk/wiki/public/solarflares/0450.html>

MOMFBD system, and they have developed a machine learning approach to correcting for these effects. This method has not been applied to the data used in this thesis.

#### 3.1.2 The Daniel K Inouye Solar Telescope

The Daniel K Inouye Solar Telescope (DKIST, [Rimmele et al. 2020](#)) is a new 4-metre Gregorian solar telescope that saw first solar light in December 2019. It is designed to support a number of instruments observing between 380 and 5000 nm with a diffraction limit of  $0.026''$  (20 km) at 500 nm. The instruments initially expected to capture flaring chromospheric observations are the Visible Tunable Filter (VTF) and the Visible Spectro-Polarimeter (ViSP). The former of these will capture narrow-band line images similar to CRISP, but at higher spatial resolution, whereas the latter will provide high-precision spectropolarimetric observations through a slit spectrograph with  $R > 180\,000$ . DKIST is expected to start making routine observations in 2021 and promises observations of the chromosphere on scales that will push our modelling and analysis techniques to their limits.

## 3.2 Introduction to Inverse Problems

From the previous chapter we have an understanding of the so-called “forward problem” of radiative transfer, that is, the synthesis of a spectrum from a known atmospheric model. The inverse of this problem is not well-posed; there is no guarantee of uniqueness, and typically the problem is extremely underdetermined. The radiation field inside the plasma couples the atomic populations at all depths (as is evident from the form of the  $\Lambda$  operator) in a way that cannot be trivially disentangled and information is then lost in the forward process that is needed for the inverse process.

It is, of course, of great value to constrain the atmospheric parameters associated with a particular observation, and it is this we seek to achieve, by solving the inverse problem of radiative transfer, known as inversion. As observed radiation is the only vector by which information can arrive from the Sun, it is important to maximally exploit the information that can be gleaned from observations and determine the structure of the atmosphere that produced the observed radiation.

The forward process is described by the diagram in [Fig. 3.2](#). Here elements of  $\mathcal{X}$  are the parameters describing an atmosphere (in this description we choose temperature  $T$ ,

3.2 Introduction to Inverse Problems

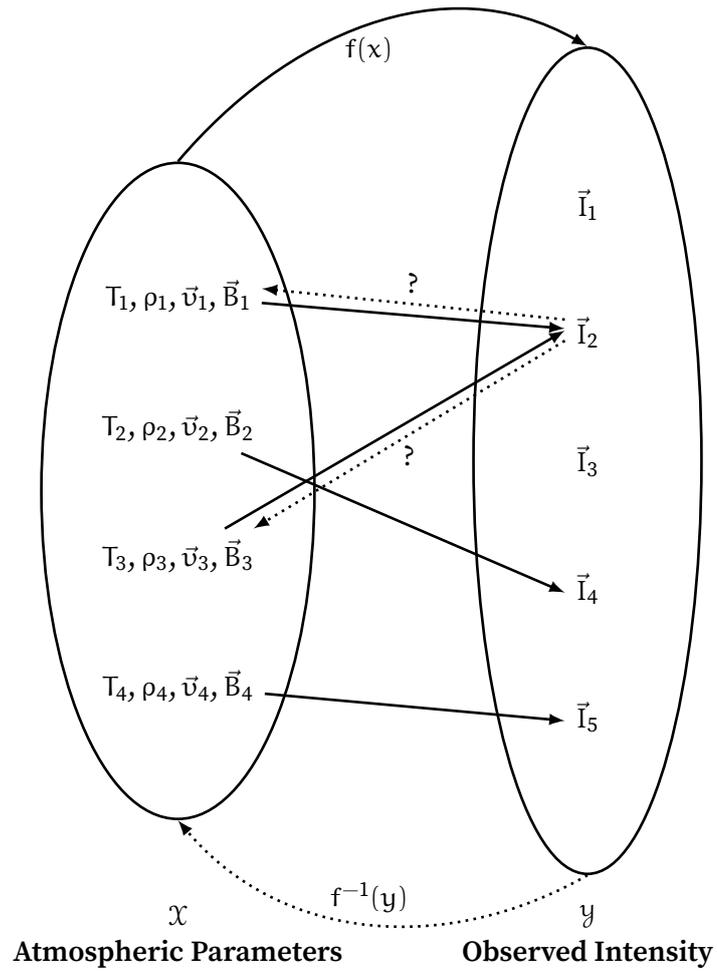
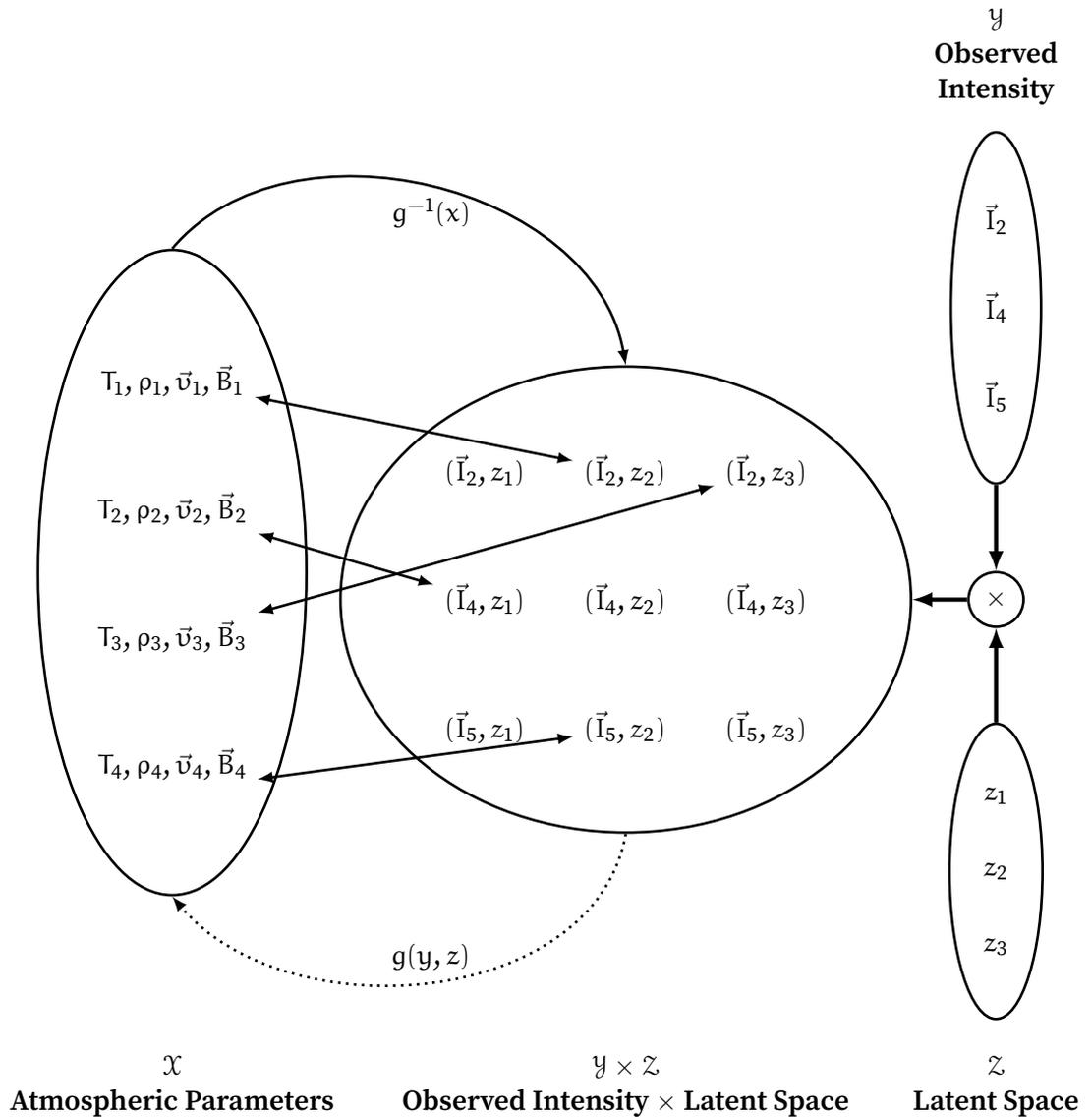


Figure 3.2: The degenerate mapping  $f$  demonstrates the difficulty of traditionally framed inversions.

3.2 Introduction to Inverse Problems



**Figure 3.3:** The bijective mapping  $g$  can be used to resolve the problem of inversions upon the introduction of a latent space  $\mathcal{Z}$  containing the information lost in the forward process. For the sake of legibility we have only included the observed intensity vectors which were mapped to in Fig. 3.2.

## 3.2 Introduction to Inverse Problems

mass density  $\rho$ , velocity  $\vec{v}$ , and magnetic field  $\vec{B}$ , although there are many other similar formulations that can be used). The elements of  $\mathcal{Y}$  are possible forms of the outgoing radiation that may be associated with different sets of atmospheric parameters. Fig. 3.2 shows the theoretically degenerate nature of the problem; whilst the mapping  $f$  from  $\mathcal{X}$  to  $\mathcal{Y}$  is always well-posed, the inverse  $f^{-1}$  is not necessarily, and with only the information present in  $\mathcal{Y}$  there is no way to immediately resolve this ambiguity (indicated by the question marks in this figure).

The problem of inversion can instead be framed as shown in Fig. 3.3, where elements of the newly introduced latent space  $\mathcal{Z}$  represent the information lost in the forward process and a bijective mapping  $g$  may be written between  $\mathcal{X}$  and the cartesian product of  $\mathcal{Y}$  and  $\mathcal{Z}$ .  $\mathcal{Z}$  is difficult to conceptualise, and harder still to characterise; in the following we will discuss multiple approaches for replacing or reconstructing this information.

As an example we can pose the simple problem of the function  $f : \mathbb{R} \rightarrow \mathbb{R}^+; x \mapsto x^2$ . This function is clearly not bijective, as for any  $y$  in  $\mathbb{R}^+$  it is ambiguous whether the expected input  $x$  was  $\sqrt{y}$  or  $-\sqrt{y}$ . If the domain of this function was instead  $\mathbb{R}^+$ , there would be no ambiguity here. Instead, we can introduce a latent space, the form of which is chosen by us and captures the lost information, i.e. the sign of the input ( $\{1, -1\}$ ). With this we can define the bijective function  $g : (\mathbb{R}^+ \times \{1, -1\}) \rightarrow \mathbb{R}; (y, z) \mapsto z\sqrt{y}$ . Discarding  $z$ , the forward process of  $g^{-1}$  is equivalent to that described by  $f$ , but the invertibility of the problem is now assured. This example does not illustrate how to discover the correct  $z$ , and for a purely mathematical case such as this, there is no unambiguous solution given only  $y$ . However, the form of the latent space is known (and in this case finite), and this can allow us to infer the different possibilities for  $x$ . In complex physical systems these different possibilities may have different likelihoods of occurring that then allow us to construct a probability distribution function for  $x$ .

For the remainder of this chapter, we denote individual samples of the atmospheric parameters  $x$ , the emergent line profiles  $y$  and the latent space  $z$ .

### 3.2.1 Milne-Eddington Inversions

*In depth reviews of the primary methods of spectral inversion are provided by del Toro Iniesta & Ruiz Cobo (2016) and de la Cruz Rodríguez & van Noort (2017).*

Milne-Eddington inversions have proven to be a simple, but very powerful tool that is often applied in solar physics. The loss of information can be somewhat limited by placing

### 3.2 Introduction to Inverse Problems

constraints on the solution. One of the simplest constraints that can be placed on the problem is that of a source function that changes linearly with continuum optical depth, whilst all other parameters are held constant throughout the atmosphere. This is a low-order approximation of the problem, and is convenient as we can express the outgoing intensity analytically. This can be done for the full Stokes polarised case of the RTE, but for illustration we choose to use only the scalar case here. With continuum opacity  $\tau_c$  the source function is then defined as

$$S(\tau_c) = S_0 + S_1\tau_c, \quad (3.1)$$

where  $S_0$  and  $S_1$  are constants. In this situation where we are dealing with the source function directly, rather than atomic parameters (as we are effectively in a two-level atom case due to only considering a single line), we can define the line strength  $\alpha$  as the ratio between the continuum opacity and the line opacity i.e.

$$\tau(\nu) = \tau_c(1 + \alpha\phi(\nu)), \quad (3.2)$$

where  $\phi$  is the line profile.

Now, by formulating  $S$  in terms of  $\tau(\nu)$  and integrating the RTE directly whilst assuming a semi-infinite atmosphere we obtain the outgoing intensity

$$I(\tau = 0, \nu) = \int_0^\infty \left( S_0 + S_1 \frac{\tau}{1 + \alpha\phi(\nu)} \right) e^{-\tau} d\tau \quad (3.3)$$

$$= S_0 + \frac{S_1}{1 + \alpha\phi(\nu)}. \quad (3.4)$$

Thanks to the analytic nature of this solution, it is easy to attempt to fit this to observations, and effects such as constant atmospheric velocity, and magnetic field can be included in the line profile. This approach is rapid and works quite well for quiet photospheric lines, such as Fe I 6301 Å and Fe I 6173 Å as used in Hinode SOT and HMI (e.g. [Centeno et al. 2014](#)), but the linear description of source function reduces its applicability to chromospheric lines.

#### 3.2.2 Generalisation Through Response Functions

It can be very limiting to impose a chosen source function, especially in regions where the temperature does not vary monotonically. A far more flexible approach is to attempt

### 3.2 Introduction to Inverse Problems

to fit a source function. From the previous discussion of formal solvers we know that for a discretised atmosphere wherein the source function follows a prescribed functional form over each interval we can write

$$I(\tau = 0) = \sum_i \int_{\tau_i}^{\tau_{i+1}} S(\tau) e^{-\tau} d\tau, \quad (3.5)$$

whereby we have once again assumed that  $I(\tau = \infty) = 0$ . If the source function is taken to be constant in each slab this becomes

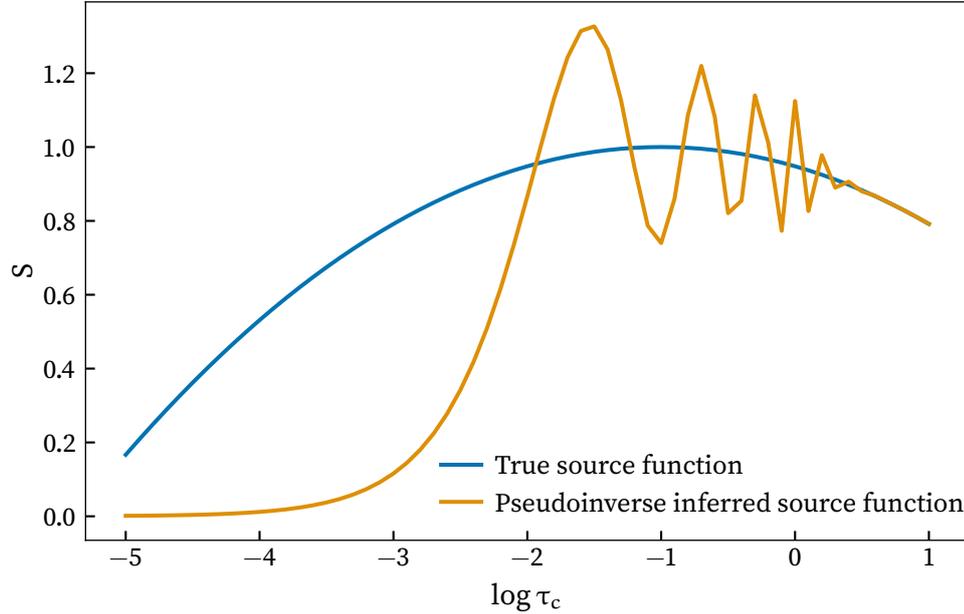
$$I(\tau = 0) = \sum_i S(\tau_i) \int_{\tau_i}^{\tau_{i+1}} e^{-\tau} d\tau, \quad (3.6)$$

and for each slab the associated integral represents its contribution to the outgoing intensity. In fact, as each contribution is linear in the source function, it represents the *response function* here i.e. the response in the outgoing intensity to a perturbation in the source function at this depth. We can therefore form a response matrix associating, for each wavelength and depth, the response to a change in source function at this depth. This response function does not change with the source function due to the linearity of the problem. In theory, we should therefore be able to infer the depth stratified form of the source function. However, two primary difficulties arise.

Typically observations of a line contain many more wavelength points than the number of depth points it is feasible to have in our discretised source function, leading to an overdetermined system with no direct inverse. Even if one were to modify the system so that these dimensions are compatible, the response matrix typically has extremely poor conditioning (is extremely sensitive to errors in the input) and thus cannot be inverted directly. An immediate solution is then to use a pseudoinverse based on the singular value decomposition of the response matrix, as this can attempt to tackle both problems at once.

Applying this method to a simple example with a quadratically varying source function stratified across 61 depth points using a constant line strength and line profile with 101 wavelength points we find the result shown in Fig. 3.4. The poor quality of the pseudoinverse solution is apparent from the large oscillations and poor agreement with the true source function, and this example showcases the simple case of a static atmosphere with constant line strength. If these parameters are also allowed to vary then the solution is likely to deteriorate further.

Whilst this problem can be overcome (to some extent) by regularisation of the solution, to



**Figure 3.4:** Comparison between true source function and inferred for the case of a quadratic variation.

enforce smoothness in the ill-posed problem, we are also left with a problem of interpretation for NLTE problems. Ultimately, we seek to learn information about the atmospheric structure from the outgoing radiation, and not simply the source function. A natural solution to this is to take a model atmosphere and attempt to fit the synthetic spectrum generated from this model to the observations, but to do this the parameters of the model need to be connected to the synthetic spectrum.

It is common to use the integral form of the RTE

$$I(\nu) = \int_0^{\infty} S(\nu, t_\nu) e^{-t_\nu} dt_\nu, \quad (3.7)$$

and define the contribution function as the integrand of this expression (Carlsson & Stein 1997; del Toro Iniesta 2003), similarly to the contribution term from (3.6). This contribution function is often interpreted as a description of the location at which a spectral line forms. If this can be used to understand how and where the line forms then it seems reasonable to use this as the basis of an inversion scheme. However, there are several problems that would arise from this approach. In many lines, the contribution at the line core can span a relatively large region, and thus there is a high probability that photons measured

at the same wavelength have been produced in significantly different locations. It is difficult to meaningfully ascribe mean thermodynamic parameters of line formation to a non-uniform extended region over which the line forms. This problem becomes more complex still if the contribution function is double-peaked, as is relatively common for lines forming in the complex atmospheres produced by flare models. As commented by [del Toro Iniesta \(2003\)](#) the mathematical definition of the contribution function  $C_I$  is also ill-posed, as for any function  $f$  such that  $\int_0^\infty f(t_\nu) dt_\nu = 0$ ,  $C_I + f$  is also an equivalently valid contribution function.

The contribution function will qualitatively capture most of the information about the regions important to the formation of a certain line, but it misses the non-local effects that occur in NLTE radiative transfer as the radiation field from one region of the model atmosphere often determines the populations and thus the emissions in another. These effects can instead be captured by the use of response functions.

Response functions have been primarily used in the field of inversions, but also represent strong tools for understanding the theory of spectral line formation in atmospheric models, due to the close coupling of these two problems. A response function tells us how the outgoing intensity changes *in response* to a change in an atmospheric parameter. These were first named by [Beckers & Milkey \(1975\)](#), but a similar concept was previously presented by [Mein \(1971\)](#) in the form of weighting functions for the RTE. They were generalised to full Stokes radiative transfer by [Landi Degl'Innocenti & Landi Degl'Innocenti \(1977\)](#) and applied by [Ruiz Cobo & del Toro Iniesta \(1992\)](#) for the situation of full Stokes synthesis and inversion (using the formulation of [Sánchez Almeida \(1992\)](#)). A similar approach (built on framing the intensity perturbations as a Fredholm integral equation and using the finite difference method to evaluate necessary terms) for the scalar RTE was employed in the inversions of [Metcalf et al. \(1990a\)](#), but their focus was on the direct application to inversions rather than the interpretation of the response functions produced.

In the following we shall consider only the response of the unpolarised Stokes I component, as we are focusing on unpolarised models, although the techniques discussed here can easily be applied to the full Stokes case. The frequency- and depth-dependent response function  $\mathcal{R}_{\nu,q}(z)$  to a parameter  $q$  can then be defined by

$$\Delta I(\nu) = \int \mathcal{R}_{\nu,q}(z) \Delta q(z) dz, \quad (3.8)$$

where  $\Delta I(\nu)$  is the change in outgoing radiation due to a change of  $\Delta q(z)$  in parameter  $q$ .

### 3.2 Introduction to Inverse Problems

To learn from response functions, we first need to know to which parameters we wish to know the intensity response. This will depend on the parametrisation of the model atmosphere used, but it is common to see response functions to temperature, electron density, velocity, and in the case of spectropolarimetry, magnetic field. An example of using response functions for chromospheric diagnostics, and their advantages over contribution functions was presented by [Uitenbroek \(2006\)](#). It was shown that there can be substantial differences between the contribution and response functions for lines that form in NLTE conditions, and highlighted that for  $H\alpha$  the photospheric conditions will affect the line-core formation height, and hence the source function due to radiative coupling.

In LTE, where the source function is set by the local atmospheric parameters, the response functions can be computed with relative ease. The SIR inversion code ([Ruiz Cobo & del Toro Iniesta 1992](#)) analytically computes the (full Stokes) response functions to perturbations in different atmospheric parameters at the same time as the formal solution. These response functions are used in conjunction with a Levenberg-Marquadt damped least squares regression procedure to modify the starting atmosphere (defined on equidistant nodes in  $\log \tau$  and assumed to follow cubic splines between nodes) until the synthesised radiation matches the observation as closely as possible.

For lines that are formed well outside LTE the process of determining the response functions to atmospheric perturbations can be significantly more arduous. The most common approach has been to apply a finite difference method to the outgoing radiation from the statistical equilibrium solution to an atmosphere by successively perturbing each parameter at each node in the atmosphere. Whilst computationally expensive, this “brute force” approach to response functions makes them simple to calculate. Following this approach, the parameter we wish to know the response to is perturbed at one depth point in our discretised atmosphere, and the populations are updated using the new radiation field and atmosphere. This procedure is repeated for the parameter at each depth in the atmosphere. The intensity response to perturbation  $\delta q$  in parameter  $q$  at depth point  $k$  can then be computed by finite differences as

$$\mathcal{R}_{\nu,q}(k) = \frac{I(\nu, q_k + \delta q) - I(\nu, q_k)}{\delta q}. \quad (3.9)$$

Despite the effective doubling in computational cost, a centred finite difference method was recommended by I. Milić (*private communication*) and [de la Cruz Rodríguez & van](#)

### 3.2 Introduction to Inverse Problems

Noort (2017), and this has proven more robust. The response function is then written as

$$\mathcal{R}_{\nu, q}(k) = \frac{I(\nu, q_k + \delta q/2) - I(\nu, q_k - \delta q/2)}{\delta q}. \quad (3.10)$$

This process must be undertaken for each depth (or node in the model) and atmospheric parameter independently, leading to its significant computational cost. In a statistical equilibrium case, the standard procedure of formal solution and population update is repeated until convergence is reached.

Whilst this process is computationally expensive, it has been used reliably since the NICOLE code (Socas-Navarro et al. 2015, developed from Socas-Navarro et al. (2000) but no longer using fixed departure coefficients). The Stockholm Inversion Code (STiC) also follows this procedure, using a modified form of RH, allowing for the application of PRD (de la Cruz Rodríguez et al. 2019).

Analytic response functions for the multi-level NLTE problem were first derived by Milić & van Noort (2017), and are now implemented in the SNAPI code (Milić & van Noort 2018). These response functions should significantly reduce the computational cost of NLTE inversions, a necessity for inverting the large fields of view in current and next generation solar observations.

An alternative approach to reducing the computational overhead of NLTE inversions can be seen in the DeSIRE code, which combines SIR and RH, guiding itself to an approximate solution using the fast analytic LTE response functions of SIR and fine-tuning the solution with finite-difference response functions computed with RH (B. Ruiz Cobo et al. *in preparation*).

All of the codes discussed here use the Levenberg-Marquadt regression method with different varieties of regularisation to enforce smooth solutions. Additionally, only the statistical equilibrium solution is considered, and then only in hydrostatic equilibrium as this reduces the number of parameters to be inferred. It is common to allow line-of-sight velocity as a parameter, which technically violates the constraint of hydrostatic equilibrium, however this is a minor effect and is seen as a worthwhile trade-off for the increase in tractability of quiet sun inversions. Clearly these constraints render this technique very difficult to apply to flares, although NICOLE has been applied to flaring atmospheres by Kuridze et al. (2018). The integral approach of (Metcalf et al. 1990a) was also applied to flares and differs from the others discussed here, by constructing a matrix of kernels which is solved for temperature and electron density perturbations using a regularised method. The kernels presented cannot capture the complete response to

perturbations in parameters affecting NLTE spectral lines without an approach that can find the new source function (which can be modified non-locally). These kernels therefore have a narrow range of validity, and require a starting guess close to the true solution.

Response function driven inversions are not the only form of inversions to have been applied, but they are the most widely used. For example, [Asensio Ramos et al. \(2007\)](#) presented a Bayesian approach utilising a Markov Chain Monte Carlo method with Milne-Eddington atmospheres. This method is too computationally costly to reasonably apply to the full depth-stratified NLTE problem with today's technology, but it would allow the investigation of the posterior distributions of the atmospheric parameters, and be less sensitive to local minima. With the recent rise in popularity of machine learning, tractable approaches to inversion that allow characterisation of the atmospheric posteriors are a key area of interest (e.g. [Osborne et al. 2019](#); [Díaz Baso et al. 2021](#)), and we will discuss these further, along with other machine learnt approaches in Chap. 7.

### 3.2.3 Forward Modelling

A technique related to inversions that has commonly been applied to flares in the last decade is that of forward modelling through the use of RHD codes. Where possible, the energy input parameters are constrained from observations, via techniques such as X-ray spectroscopy to deduce the non-thermal electron flux and spectral index. A challenging manual iteration then follows to attempt to obtain an agreement between the time-dependent simulation and the observations. This is extremely time-consuming both due to the manual aspect of the inversions and the computational requirements of the RHD simulations. Clearly, the human intervention necessary to optimise and analyse these simulations cannot scale to the large volumes of data coming already present and coming from future telescopes. This technique has nevertheless yielded many interesting developments in our understanding of the structure of the flaring chromosphere from the investigation of both spectral line shapes and continuum variations ([Kuridze et al. 2015](#); [Rubio da Costa et al. 2016](#); [Kowalski et al. 2017b](#); [Simões et al. 2017](#)).

### 3.2.4 In the Context of the Latent Space

Returning now to the mathematical description of an inversion we can start to discuss the meaning of  $\mathcal{Z}$  in practice. With the response function based inversions described

### 3.3 Introduction to Machine Learning

previously, the size of  $\mathcal{Z}$  is limited by the constraints placed on the atmospheric stratification, and the regularisation thereof. It then becomes feasible to “explore” this space ( $\mathcal{Z}$  coupled with  $\mathcal{Y}$  as no explicit distinction is made) using the gradient information from the response functions to guide the solution. It is worth noting that this approach does not guarantee the global minimum solution; whilst the Levenberg-Marquadt algorithm is extremely efficient at finding local minima, it provides no further guarantees and the final solution may therefore be substantially influenced by the choice of starting atmosphere, which is typically picked based on intuition from the “standard” semi-empirical models. The RHD based forward fitting methods are also comparable in terms of exploration of  $\mathcal{Z}$ , except here the optimisation is done manually and the gradient information is replaced by intuition.

### 3.3 Introduction to Machine Learning

Machine learning describes a family of generic algorithms that are used to make sense of data without being explicitly programmed. A model is defined by the researcher, but its final behaviour is determined by patterns in the data it is fed. The abundance of both observational and simulated solar data continues to increase and new approaches, such as machine learning, are needed to make use of this vast quantity of information in a computationally tractable manner, helping to highlight patterns that can be further investigated by researchers.

There are three primary varieties of machine learning algorithms: supervised, semi-supervised, and unsupervised learning. Supervised algorithms are the most common. The model is provided with a set of examples (typically produced or preprocessed manually) and is then trained so that it represents an approximate transformation between the input and output data defined by the training data. We can further divide this class into classification and regression models. Classification associates each class with a discrete input, possibly labelling an image based on its contents, whereas regression approximates a continuous mathematical function. In both of these cases the model approximates a function which is learnt entirely from the training data.

Unsupervised learning does not require the manually prepared set of examples, but instead organises data based on generic programmed criteria. Two commonly used examples of unsupervised learning are clustering and dimensionality reduction techniques. Clustering algorithms extract groups of similar objects (where similar is defined given a particular

### 3.3 Introduction to Machine Learning

basis and metric determined by the choice of algorithm), and can be used to find patterns in large datasets. There are many kinds of dimensionality reduction techniques, but one of the most common and general choices is principal component analysis, where an orthogonal basis spanning the data is constructed and then sorted by the variance of the factors of each of these axes (i.e. the eigenvalues of the covariance matrix). For data of dimensionality  $m$ , keeping  $m$  principal components allows for a perfect reconstruction, as this is simply a basis transformation, however, we can often discard terms with small variance and produce accurate approximate reconstructions of the data with substantially fewer than  $m$  components. It is necessary to ensure that sufficient components are chosen for the reconstruction to be accurate, but such techniques can reveal patterns that are otherwise difficult to discern in the original high-dimensional spaces.

Finally, as implied by the name, semi-supervised learning lies in between the two previously discussed classes. It still requires preprocessed training data which is used for some training, however unsupervised learning processes may be used internally to the model, or in some cases data generated by a model is used in conjunction with this training data. This form of machine learning exists only within the realm of deep learning, built on neural networks.

#### 3.3.1 Artificial Neural Networks

Artificial Neural Networks (ANNs) loosely follow the principle of biological neuronal systems, consisting of layers of interconnected neurons, the output of which is summed in synapses and then has a non-linear activation function applied to determine if the signal is passed on through the network. ANNs consists of multiple layers of neurons and synapses whereby we designate any layer that is neither the input nor the output a *hidden layer*. If an ANN consists of more than one hidden layer it is termed a deep neural network (DNN), and these are considered to be the standard building blocks of modern machine learning (Raschka 2015).

There are many different architectures for ANNs, used for solving different problems. ANNs can vary in number of hidden layers, interconnectedness of neurons within these layers, connectedness of the layers to each other, and the activation function used in each layer. We distinguish two primary forms of layers, based on their interconnectivity; these are fully connected (FC) where each neuron in a layer is the linear combination of its inputs (typically the activation function applied to the neurons of the previous layer), and the convolutional layers of convolutional neural networks (CNNs, Lecun et al. (1998);

### 3.3 Introduction to Machine Learning

Simard et al. (2003)) which connect only nearby neurons to exploit local structure in the input (in one or more dimensions). These convolutional layers can then be described as a set of filters learned during the training process which are cross-correlated with the input, the output of which is then passed through the activation function. CNNs are somewhat inspired by the neuronal structure of the visual cortex, and have been applied with great success in the fields of image analysis, processing, and generation (Raschka 2015). A fully connected layer rarely works well for these tasks as a slight movement of an object within an image can easily invalidate its training whereas the smaller layers of a CNN sweep across the image (are applied to each region in turn) and are far less affected by this.

There are many common forms of activation function. Due to the backpropagation method used for training ANNs it is highly advantageous if these non-linear activation functions be trivially differentiable. Some common choices are the sigmoid function

$$S(x) = \frac{1}{1 + e^{-x}}, \quad (3.11)$$

inverse tangent  $\tan^{-1}(x)$ , and variants of the rectified linear unit (ReLU; Nair & Hinton (2010)).

$$\text{ReLU}(x) = \max(0, x). \quad (3.12)$$

All of these functions are used in the creation of ANN based models, but the ReLU family is key to modern machine learning for reasons of sparsity in its output. Here sparsity refers to the presence of zeros in the output of a layer creating clearer pathways through the network. Classification ANNs will typically employ an activation function on the output layer (most frequently a normalised exponential to select a single discrete class), whereas ANNs employed in regression problems will rarely do so.

#### 3.3.2 General Function Approximations

ANNs are universal function approximators; they can learn arbitrarily complex classification and regression problems (Rumelhart et al. 1986; Cybenko 1989). This was theoretically proven for shallow neural networks (with only one hidden layer) using sigmoidal activation functions by Cybenko (1989). Increasing the precision to which a function is approximated may however require exponential increases in layer width and training. A similar proof for the commonly used ReLU activation function was provided by Lu et al. (2017), who developed bounded expressions for the layer width and network depth needed to approximate functions to arbitrary precision. Unfortunately these results can be difficult

### 3.3 Introduction to Machine Learning

to apply to many real world scenarios where the intrinsic dimensionality of the function being approximated is not known (these results are also affected by any imperfections in the training data).

It is also possible to increase the approximation capability of an ANN by increasing its depth (the number of stacked layers), these stacked layers then represent the composition of functions, and each additional layer increases the complexity of the representation of its input, allowing for very complex tasks to be approximated (Raschka 2015). The approximation power of stacked layers explains why the DNN is core to modern machine learning, however care must be taken when designing a model to select appropriate width and depth for the problem at hand (Lu et al. 2017).

#### 3.3.3 Training via backpropagation

ANNs are trained via a process known as backpropagation (Rumelhart et al. 1986). The networks are composed of linear combinations and (by our original requirements) differentiable activation functions. The entire network can then be differentiated by repeated applications of the chain rule (from output to input) to find the gradients of the output with respect to each weight (the coefficients of the linear combinations in each layer) and input, which then describes how each weight affects the output. Typically the output of the network when fed with data from the training set is compared against the expected output via a loss function, and then the gradient information from this loss is combined with the previous gradient of the network output to each weight and used to minimise the magnitude of the loss by modifying the weights.

Updating the weights in the network can be carried out in a variety of ways, but it is a similar minimisation process to that used in inversions. The basic method is that of stochastic gradient descent (SGD) which takes a step through the loss space guided by the gradients for each batch of training data. The size of this step is known as the learning rate, and is a *hyperparameter*<sup>3</sup> of the ANN. It can be kept constant, vary following a prescribed evolution with epoch, or even be modified based on the rate of convergence of the training procedure. As SGD is only affected by the most recent batch of data it can have difficulty escaping local minima and traversing plateaus in the loss space.

Many improvements to SGD have been developed, such as the addition of momentum, which accelerates convergence and helps to avoid the solution being overly affected by

---

<sup>3</sup>Hyperparameters are tunable parameters that are often set by the researcher, or optimised by a process external to the training of the INN

### 3.3 Introduction to Machine Learning

a single batch of training data. Other modern algorithms based on the same principles as SGD have also been developed (e.g. the Adam algorithm, [Kingma & Ba 2014](#)) and often converge in fewer epochs (rounds of training) to similar or better solutions. None of these stochastic algorithms can guarantee a global minimum in the loss space, and such a requirement is not feasible for anything other than the smallest neural networks, where more time- and memory-consuming optimisers can be used due to the much more dimensionally compact spaces over which the optimisation occurs. Nevertheless, with sufficient training data and epochs a model capable of approximating the function we wish to learn should be able to descend into sufficiently good local minimum using these techniques.

Auxiliary techniques to improve model convergence have also been developed, such as minibatching, in which the network is only shown a random portion of the training data each epoch. Clearly this can reduce the computational cost of an epoch, as fewer calculations are performed on this training set, but minibatching can also improve the convergence by avoiding the stagnation that arises in the traditional batched gradient descent where the entire training set is used to direct the step.

A technique known as autodifferentiation has become prominent in the field of machine learning. It allows users to easily design custom blocks and compose these without the need to consider the implementation of the derivatives needed for training as these are computed by the framework. Frameworks (e.g. TensorFlow ([Abadi et al. 2016](#)), PyTorch ([Paszke et al. 2019](#))) may record the path of data through a network and then using this information (as every function present therein is differentiable) construct all necessary gradient information for training, which can then be computed on GPU. The automatic nature of this approach has enabled the rate of development seen in machine learning in the last decade, as it allows researchers to spend longer thinking about design than low-level engineering.

#### 3.3.4 Difficulties training DNNs

As the number of layers in an ANN increases the networks can become much harder to train; the gradient of the output with respect to the weights in early layers can easily become vanishingly small due to the repeated multiplication of small gradients in the deeper layers. The use of ReLU activation functions often minimises this effect, but can instead lead to exploding gradients due to their high dynamic range. [He et al. \(2015\)](#) developed residual networks (ResNets), which have greatly increased the depth and complexity of

### 3.3 Introduction to Machine Learning

networks that can be effectively trained, and now networks with many hundreds of layers are frequently used (e.g. [Jegou et al. 2017](#)). The residual blocks of these networks contain so-called skip connections, which take the output from a layer and sum or concatenate it with the output of layer one or more levels deeper. These skip connections provide a path for gradients to propagate through the network, helping to avoid both vanishing and exploding gradients. Variants of the ReLU function are almost uniquely used in ResNets as these additionally provide sparsity to the representation (i.e. their output is 0 for all input less than or equal to 0), which can improve the expressiveness<sup>4</sup> of the representation and aid in disentangling information propagating through the network ([Glorot et al. 2011](#)). These variants include the leaky ReLU ( $\max(0.01x, x)$ , [Maas et al. 2013](#)) which still produces a small amount of gradient information for negative inputs, helping to prevent neurons with ReLU activation from “dying”, and exponential linear units (ELUs, [Clevert et al. 2015](#)) which achieve a similar result in a smoothly varying fashion.

Like all regression models with a large number of free parameters, ANNs can very easily enter a regime of overfitting their training set. In this situation the ANN has learnt to match its training set so closely that it is unlikely to perform reliably on inference of unseen data. This can often manifest as memorisation, where the network has learnt to produce the expected output for a training sample, but not the relationship between the two. ANNs must therefore be trained with care and diligent use of validation data, prepared in the same way as the training set, but never shown to the network during training. The network’s performance can be judged by how well it performs in inference on the validation set in between training epochs. If the performance on the training data continues to improve over time, but the performance on the validation set stagnates or worsens then the network has entered an overfitting regime.

There are additional techniques that can be employed to mitigate overfitting, such as regularisation, which will attempt to prevent a model’s weights from minimising the loss function too perfectly, for example by penalising overly large weights with a modified loss function, or randomly deactivating neurons in each layer during training (this approach is known as *dropout*).

Selecting hyperparameters for a model can be a challenging process of manual optimisation but is essential to training, and many advanced optimisers like Adam require additional hyperparameters that can drastically influence the rate of convergence. Approaches such as grid searches can be applied here, but given the computational requirements of training these models, an intuitive approach is often applied.

---

<sup>4</sup>The expressive power of a neural network is its ability to approximate functions.

# 4

## The *Lightweaver* Radiative Transfer Framework

The *Lightweaver* framework<sup>1</sup> (Osborne & Milić 2021; Osborne 2021a) is a Python package built around a C++ core in which we have implemented the methods for numerical NLTE radiative transfer discussed in Sec. 2.4. As can be seen from the referencing of this section, none of these methods are novel on their own – they represent the most robust methods encountered in our survey of NLTE radiative transfer – and it is in the combination of these methods and the implementation strategies employed that *Lightweaver* differs from current state-of-the-art radiative transfer codes. An overview of the key components and functions that users will typically interact with is presented in Osborne & Milić (2021), in the following we describe the most important design decisions made in *Lightweaver* and explain how they can enable new forms of radiative transfer simulations whilst also increasing productivity.

### 4.1 Philosophy

The design of the *Lightweaver* framework is inspired by deep learning frameworks, such as PyTorch (Paszke et al. 2019). These have risen to prominence in recent years, due to their low barrier to entry, whilst still providing a customisable, full-featured, interface to the underlying methods that can be manipulated with pure Python code. Machine learning frameworks provide a collection of building blocks that can be combined in multiple ways to allow researchers to construct new tools, specifically tailored to the problem they wish to address. Whilst there can be slight performance gains from using a specialised,

---

<sup>1</sup>*Lightweaver* is freely available under the permissive MIT license on GitHub (<https://github.com/Goobley/Lightweaver>) with archival on Zenodo.

## 4.1 Philosophy

optimised, state-of-the-art method implemented in a performance-focused language for this particular task, the benefits are likely outweighed by the additional development time. This is especially true in research environments where tools are often used by a small group of researchers in a transient fashion, and the return on *possible* optimisations is rarely sufficiently large compared to the benefits of increased development speed that a framework allows. The use of a tested framework also allows researchers confidence in the core numerics they are reusing, whether they understand every detail or not.

The steps involved in solving the NLTE radiative transfer problem, e.g. formal solution, calculation of preconditioned rates, population updates, conservation of charge, and calculation of the PRD line profile ratio, are quite modular, and we provide optimised methods for the most commonly used steps following the standard techniques outlined previously. They are building blocks that can be combined in different ways, to produce different tools. These building blocks are the core offering of the *Lightweaver* framework, and are intended to be combined by the user in a new Python program to solve their particular problem. If at any point a user wishes to fully replace a core component of *Lightweaver*, this can be done in Python, inside their program, with no modification of the framework itself, whilst the other components can continue to be used as before. This flexibility is encompassed by one of the core design goals, which is to allow Python code written by the user to “interfere” with all of the numerical treatment of the NLTE problem.

All other radiative transfer codes that the author has interacted with have been designed with a strict limit of one simulation per computer process. Whilst this limitation does make the design of the program easier, especially in Fortran and C, it is not beneficial to an end user who may, for example, wish to couple multiple simulations with different atomic configurations, or, say, use one as a radiative boundary condition for another. This latter configuration is applied extensively in Chap. 6 where plane-parallel models are used as boundary conditions for a two-dimensional slab. Whilst there are other solutions to this problem, such as saving necessary data and loading it in a reconfigured program, these are typically more error-prone than a simple program which can flexibly represent the coupling between these models in its code, even allowing for memory sharing of certain components. To this end, each radiative transfer simulation performed in *Lightweaver* occurs in a self-contained `Context`, a Python object containing all necessary configuration and storage needed for this model. These can be serialised using the `pickle` package of the Python standard library, allowing for a complete simulation to be dumped to disk or transferred between processes using the standard approach expected in the Python ecosystem. This has impacts on parallelisation, which will be discussed in Sec. 4.5.

## 4.2 Accessibility & Code Overview

One of the aims of *Lightweaver* is to attempt to reduce the barrier to entry for new users, thus we ensure that it is simple to install with pre-compiled libraries available. Thus a user with a Python environment (version  $\geq 3.8$ ) using an x86-64 CPU supporting AVX vector extensions (essentially any Intel or AMD chip from the past decade) on any of macOS<sup>2</sup>, Windows, or Linux, can install the package in one command using the Python package manager `pip`. No additional compilation steps are necessary, and any additional libraries required are automatically sourced during installation. Whilst slight performance benefits can likely be acquired by using a modern compiler to tune the code generated to the user's machine (and this option is available for advanced users), the option of automatically installing a tested release version of the library in under 60 s was an important goal that has easily been achieved thanks to Python's well-supported packaging systems.

All interfaces to the framework are thoroughly documented through the Python docstring convention (internally to the source files) and can be used to automatically generate HTML or L<sup>A</sup>T<sub>E</sub>X documentation. This can also be viewed online at <https://goobley.github.io/Lightweaver>.

As of v0.7.3, excluding automatically generated code (of which we make extensive use), the *Lightweaver* frontend consists of 4520 lines of Python, with 2326 lines of comments and documentation. 777 of these consist of an implementation of the equation of state originally authored by Wittmann following Mihalas (1978), and ported to Python by J. de la Cruz Rodriguez (used here with permission). This is an LTE equation of state that has been used in both the SIR (Ruiz Cobo & del Toro Iniesta 1992) and NICOLE (Socas-Navarro et al. 2015) codes.

The backend consists of 9896 code lines of personally authored C++, along with two external libraries: Faddeeva<sup>3</sup> (Steven G. Johnson, 2066 lines of code, MIT license) used for computing Voigt functions, and a lightweight multi-platform thread pool and scheduler<sup>4</sup> (Doug Binks & Micha Mettke, 788 lines of code, zlib license). As both of these libraries are small and permissively licensed, they are included directly in *Lightweaver*'s distribution, so there is no concern about these links going stale. Multiple routines present in the calculation of the background terms are thread-safe reimplementations of those used in RH (Uitenbroek 2001), with permission. There are also 2439 lines of Cython (Behnel

<sup>2</sup>Preliminary support of Apple's ARM CPUs is present, but several underlying libraries do not yet support this.

<sup>3</sup>[http://ab-initio.mit.edu/wiki/index.php/Faddeeva\\_Package](http://ab-initio.mit.edu/wiki/index.php/Faddeeva_Package)

<sup>4</sup><https://github.com/vurtun/lib/blob/master/sched.h>

et al. 2011) code present in the backend. Cython is a compiled language used to bridge the Python interfaces to the C++ core. It allows us to share NumPy (Harris et al. 2020) arrays by reference between Python and C++, allowing changes to the array’s contents to be visible from either language with no duplication necessary. This data sharing is not just efficient but allows the Python frontend to be “involved” with the radiative transfer calculations on a deep level in line with *Lightweaver*’s design goals.

## 4.3 Model Atoms

An oft-quoted aphorism in programming circles is that of *Greenspun’s tenth rule of programming*<sup>5</sup> which states: “Any sufficiently complicated C or Fortran program contains an ad hoc, informally-specified, bug-ridden, slow implementation of half of Common Lisp”. The implementations of model atoms in the codes the author is personally familiar with are examples of this. This is not to say that configuration files containing the data needed to run a program are problematic, but we are instead referring to the large amount of logic associated with these files, that eventually turns into an *ad hoc* domain specific language. These models are structured and contain methods of specifying the approximations to be used for different terms, such as van der Waals broadening and bound-free cross sections. Any new method that a user wishes to implement then has to be added to the custom interpreter responsible for parsing these files and propagating this information into the numeric core of the program.

A different approach to the problem of needing to specify data with *associated* methods and approximations is to define a set of requirements (henceforth *contract*) specifying the information needed by the numerical core and allow the specific model to fulfil this contract through any means. It is this approach we take in *Lightweaver*. This means that model atoms need to be “smart” and have the ability to execute arbitrary code. Implementing these models in Python makes this trivial, and its wide array of scientific libraries are also available<sup>6</sup>. As a “free bonus”, thanks to the models being standard Python objects stored in source code, the Python interpreter will take care of parsing these models, through extremely well-tested code paths.

---

<sup>5</sup>e.g. <https://philip.greenspun.com/research/>

<sup>6</sup>We stress that other languages could be used for this task, and they need not be dynamic “scripting” languages; for example, a similar approach could be achieved in C/C++ through the use of dynamic libraries.

#### 4.4 Other Implementation Details

An `AtomicModel` in *Lightweaver* is the definition of an object containing an element or isotope identifier, and a list of each of levels, lines, continua, and collisional rate approximations. Each of these terms is itself a Python object which must conform to a particular contract describing the form of the information it must be capable of supplying when a particular method is called. Many of these objects also contain components with their own contracts. These contracts are defined through the use of Python classes, and a basic implementation of features, comparable to those in extant radiative transfer codes, is present within the core of *Lightweaver*. These can be further extended with new functionality through the use of inheritance. A user can therefore implement a new method for computing, say, the absorption profile (e.g. the non-Voigt profile of [Kowalski et al. 2017b](#)) of a spectral line and employ this by defining a new model atom with no changes needed in the *Lightweaver* package.

Atomic models can be supplied in two different ways, in textual source code, which is both human and machine readable, or in a `pickle`, which is only machine readable, possibly from a previously serialised `Context`. The former of these is treated as a canonical form, and it can always be recovered by asking Python for the representation of the object (via the `repr` function). This imposes a minor constraint on the way user constructed classes extending the components of these models need to be defined i.e. they must define a `__repr__` function defining how to recover a textual representation of themselves. This is due to our adherence to the standard Python convention that `obj == eval(repr(obj))`, requiring that the evaluation (in the Python interpreter) of the textual representation of an object give an equivalent object. This is explained in depth with the examples provided in [Osborne & Milić \(2021\)](#).

The use of Python objects and data structures to describe the model atoms does not only allow the models to execute arbitrary code, but also for the models to be manipulated via user code. This makes managing and modifying atomic models easier as transformations (modifications of parameters) can be undertaken in bulk by code, rather than a painstaking manual process. This is all achieved through the use of standard Python, with no custom code needed to support this.

#### 4.4 Other Implementation Details

Here we present a general overview of the implementation of some of the techniques discussed in [Sec. 2.4](#), primarily from a design and usage perspective, rather than a nu-

#### 4.4 Other Implementation Details

merical one, as we find the former to be more insightful. The numerical implementation details of *Lightweaver*, including ensuring correct normalisation of the numerical integrations performed are presented in depth in Osborne & Milić (2021). In general we follow the best practices discussed in literature, and given the amount of memory present in today's computers we typically opt for slightly more memory costly approaches if they are significantly simpler or more efficient. In the plane-parallel case, a Gauss-Legendre quadrature based on the polar angle of the rays is typically used for integrating terms with angular dependence, although a user can supply their own that is optimised to a particular problem. Integrations over frequency are typically carried out by using a simple Riemann sum variant, such as the midpoint rule, and ensuring normalisation of terms such as the line absorption profile and PRD scattering integral.

In *Lightweaver*, model atoms are considered *active*, *detailed static*, or *passive*. Prior to construction of the `Context`, the user defines the set of model atoms, and how each is to be treated in this particular simulation. This is done through an instance of the `RadiativeSet` class. Active atoms are given the full NLTE treatment, with the terms necessary for iterating the populations (the preconditioned  $\Gamma$  matrix) being computed with the formal solution. To correctly resolve the effects of Doppler shifts, the emissivity and opacity terms associated with the transitions of these atoms are considered to be angle-dependent, and are computed for each ray in the angular quadrature, both up- and downgoing. So-called detailed static atoms are treated similarly, but the terms for updating the populations are not computed. As such, they will be ignored by the population update routines. A model atom defined as detailed static is intended to be used when the atomic populations of a species are known *a priori*, e.g., it is in LTE, or its populations determined in an alternative manner but its transitions are a significant source of emissivity and opacity that may affect another species being simulated (and the radiative terms associated with this other species are not expected to have a significant effect on the former). Model atoms designated as passive are treated in LTE and only their bound-free transitions are considered. Their contributions to the plasma emissivity and opacity are considered isotropic and incorporated into the background terms. The populations of all of these species are stored in a `SpeciesStateTable`, which holds the arrays that are shared with the C++ iteration machinery.

To handle overlapping transitions, we treat all of these on a common grid. This grid is computed by the instance of `RadiativeSet`, after the treatment of each atom has been defined. The wavelength grids of each transition considered in detail are combined, and the region of this common grid associated with each transition is used in the evaluation of its line absorption profile or photoionisation cross-section. During the formal solution, all

#### 4.4 Other Implementation Details

of the sources of emissivity and opacity present at each wavelength (active, detailed static, and passive model atoms, along with other background terms) are summed to provide the total emissivity and opacity terms needed to compute the source function.

As discussed in Sec. 4.2, the background emissivity, opacity, and scattering terms take the same form as those in RH. The background implementation is defined through a flexible interface that receives the atomic populations, model atoms, and wavelength at which the background parameters need to be known. Through this interface, a user can easily override the provided implementation. In all of the numerical experiments presented in this thesis the default implementation (or its parallel cousin `FastBackground`) is used. The components considered in the default background implementation and their original references are:

- Frequency independent Thomson scattering (Mihalas 1978)
- H free-free (Mihalas 1978)
- $\text{H}_2^-$  free-free (Bell 1980)
- $\text{H}_2^+$  free-free (Bates 1952)
- $\text{H}_2$  Rayleigh scattering (Victor & Dalgarno 1969; Tarafdar & Vardya 1973)
- $\text{H}^-$  bound-free (Geltman 1962; Mihalas 1978)
- $\text{H}^-$  free-free (Stilley & Callaway 1970; Mihalas 1978)
- $\text{H}^-$  free-free ( $> 9113$  nm) (John 1988)
- basic Rayleigh scattering for H and He (Mihalas 1978)
- OH bound-free (Kurucz et al. 1987)
- CH bound-free (Kurucz et al. 1987)
- Bound-free terms from all passive atoms.

We note that the terms with (Mihalas 1978) as their reference are also discussed similarly in Hubený & Mihalas (2014). Due to its importance in obtaining a correct background opacity, the molecular  $\text{H}^-$  population is always computed. *Lightweaver* can compute the formation of molecules, assuming that they form in instantaneous chemical equilibrium. Their formation will reduce the total populations of the species bound up in them, and they may contribute to the background opacities (e.g. CH, OH,  $\text{H}_2$ ).

In *Lightweaver*, the time-dependent population updates are computed using the fully implicit variant of the method described in Sec. 2.4.11, i.e. with  $\theta = 1$ . In general, we have found no difference between the fully- and semi-implicit treatments of the term, and adopt  $\theta = 1$  for simplicity. It is simple to replace the time-dependent population update

## 4.5 Parallelisation

scheme with one that performs the semi-implicit integration without modifying the core of *Lightweaver* itself.

The derivatives of the collisional rates needed for the Newton-Raphson charge conservation scheme are computed by a finite-difference approximation. This is relatively efficient as these rates depend only on local parameters, and much less restrictive than requiring a user to also compute the analytic derivative of any collisional rate formalisms they add to their own model atoms.

As discussed in Sec. 2.4.2, we adopt a short-characteristics formal solver to solve the RTE, and in *Lightweaver*, the default implementation for plane-parallel atmospheres is the cubic Bézier spline technique of [de la Cruz Rodríguez & Piskunov \(2013\)](#). We also provide an implementation of the simple linear short-characteristics scheme, and allow for users to load their own formal solvers from shared libraries at runtime, provided they implement the interface described in the *Lightweaver* codebase.

## 4.5 Parallelisation

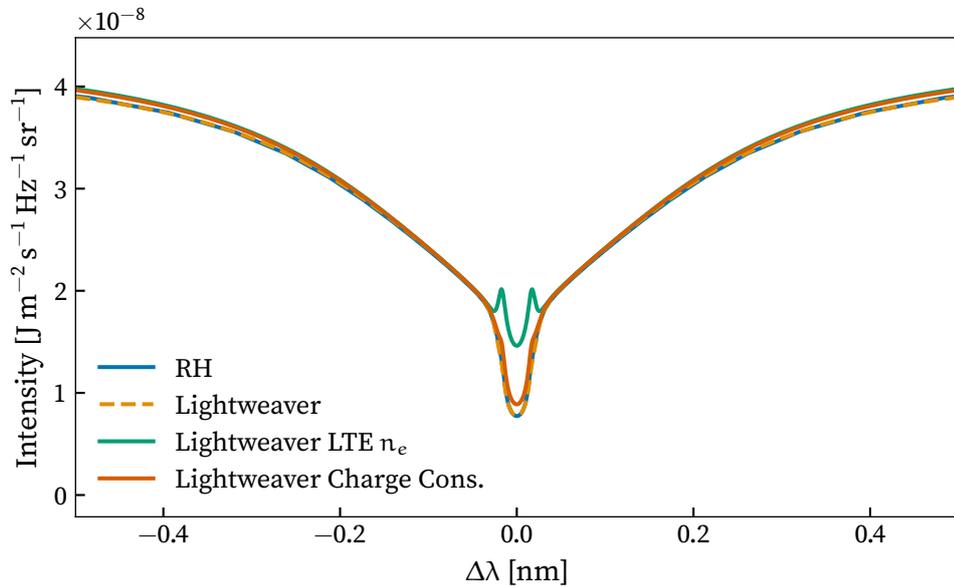
The self-contained nature of the `Context` makes *Lightweaver* programs for computing grids of models easy to adapt to paradigms such as the Message Passing Interface (MPI, e.g. [Gropp et al. 1996](#), for an overview of the MPICH implementation) commonly used in high performance computing environments. A proof of concept implementation utilising MPI for a grid of models was undertaken by A. Asensio Ramos (*private communication*), running 10,000 models in 4 hours across 15 CPUs with no need to modify *Lightweaver*.

A secondary form of parallelisation is also incorporated into *Lightweaver*. This consists of splitting time-consuming work from a single simulation over multiple threads in the same machine. This work primarily consists of the formal solution, accumulation of terms into the  $\Gamma$  matrix, and calculation of any line profile ratios needed for PRD, which are parallelised over wavelength. These terms typically represent the vast majority of the program's runtime for non-trivial simulations. The calculation of line absorption profiles, when using atomic models with the default Voigt profile implementation, along with the optional `FastBackground` implementation of background opacities, emissivities, and scatterings, are also parallelised. The choices of the terms to parallelise has been motivated by our observations of the most time-consuming processes when using *Lightweaver* to undertake the numerical experiments presented in Chaps. 5 and 6.

## 4.6 Validation

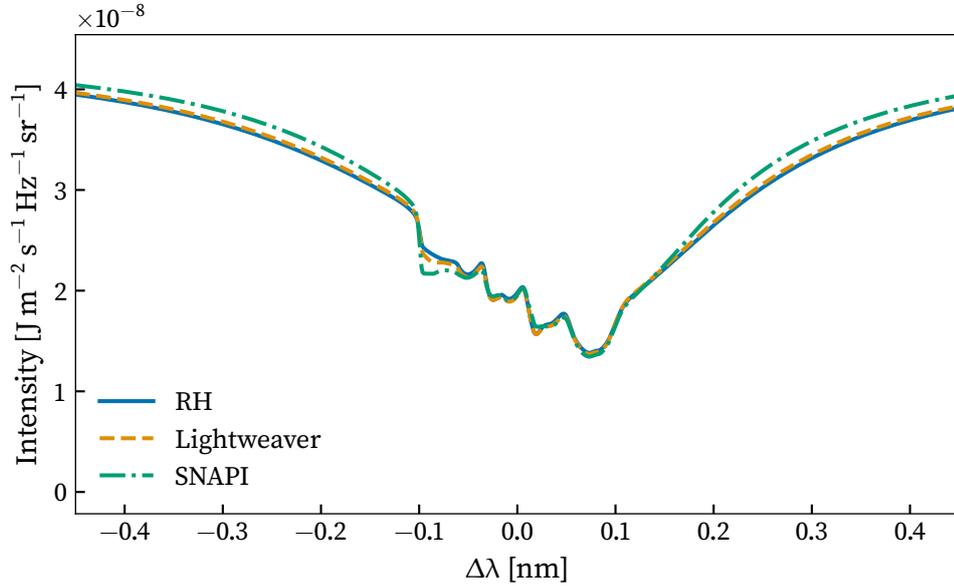
The *Lightweaver* framework was extensively validated during development, primarily against RH (Uitenbroek 2001), but also the synthesis module of SNAPI (Milić & van Noort 2018) when discrepancies were found. RH is a well-established code that serves as a cornerstone of NLTE radiative transfer in the solar physics community. It assumes a single time-independent atmospheric input, for which the statistical equilibrium solution of the atomic populations is computed. The MALI method with full preconditioning (Rybicki & Hummer 1992) is used, and implementations of angle-averaged and angle-dependent PRD with cross-redistribution are present following the methods outlined in Uitenbroek (2001) and Miller-Ricci & Uitenbroek (2002). RH can also be used for NLTE modelling of molecular lines, but we do not make use of this here. In the examples presented here, we make use of v2 of the RH code, distributed by H. Uitenbroek, and not the massively parallel RH 1.5D presented in Pereira & Uitenbroek (2015).

The synthesis module of SNAPI also uses a MALI method allowing for overlapping lines, with an implementation of the method entirely distinct to the one used in RH. It assumes a CRD treatment of all spectral lines. In the following we will present just a few of the validation cases that have been used during the development of *Lightweaver*.



**Figure 4.1:** Comparison of *Lightweaver* and RH synthesis of Ca II 854.2 nm from the FALC atmosphere with different electron density solutions.

## 4.6 Validation



**Figure 4.2:** Comparison of the *Lightweaver*, RH, and SNAPI synthesis of Ca II 854.2 nm from the FALC atmosphere with complex velocity profile and LTE electron density.

Fig. 4.1 presents a comparison of the synthesis of Ca II 854.2 nm by *Lightweaver* and RH in the FALC atmosphere of Fontenla et al. (1993). This is a static, semi-empirical quiet Sun atmosphere, so the synthesis is performed in statistical equilibrium. RH’s solution is shown in blue, with the *Lightweaver* solution overlaid in dashed orange. There is clearly very good agreement between these codes in this simple static test. This figure also shows the sensitivity of Ca II 854.2 nm to the electron density when computed under the assumption of LTE ionisation (green), and how the charge conservation method implemented in *Lightweaver* can help to mitigate these effects when the electron density is not known *a priori* (red). The line profile computed from the simulation with the charge conservation strategy approaches the reference solution, where the electron density is provided by the FAL model. The primary differences between these treatments is likely due to other species, such as Fe, being treated in LTE.

In Fig. 4.2 we once again present the synthesis of Ca II 854.2 nm in the FALC atmosphere, but here a complex velocity profile is imposed on the atmospheric model. The vertical velocity structure of the atmosphere is defined to be two periods of a sine wave with amplitude  $30 \text{ km s}^{-1}$  mapped over the vertical extent of the FALC model (approximately 2 Mm). The electron density is set and held fixed at the value given by the LTE ionisation state of the plasma. This is to allow for easier comparison between the codes with different

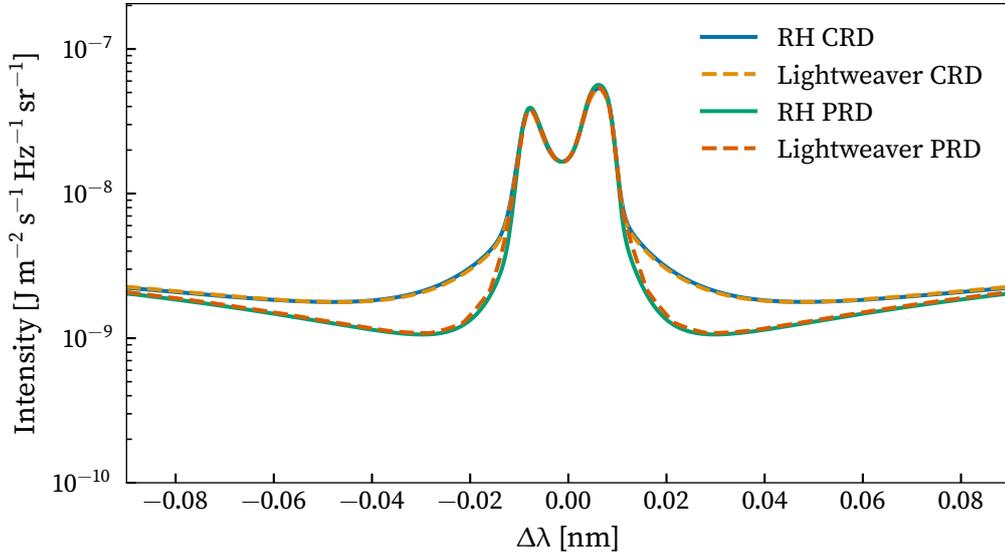
## 4.6 Validation

formulations for their model atmospheres. Here RH's solution is once again shown in blue, *Lightweaver's* in dashed orange, and SNAPI's in dot-dashed green. This problem serves primarily as a test of the formal solver: the piecewise parabolic method of Auer & Paletou (1994) implemented in RH, which falls back to monotonic linear interpolation if the parabolic terms under- or overshoot, was unable to correctly solve the problem, causing the iteration to fail by producing a singular  $\Gamma$  matrix. The cubic Bézier method of de la Cruz Rodríguez & Piskunov (2013) normally used in all three of these codes, along with their standard MALI iteration machinery, provides a very similar solution in all three cases. The slight differences apparent around  $\Delta\lambda = -0.1$  nm is likely to be caused by differences in the formal solver implementations; the version implemented in RH uses the method of Fritsch & Butland (1984) to compute numerical estimates of the necessary derivatives, whereas we use the method of Steffen (1990) in *Lightweaver* following the recommendation of Janett et al. (2018), as the accuracy of Fritsch & Butland (1984) falls to first-order on non-uniform grids. Additionally, the implementation present in RH chooses to limit the control points on the spline interpolants to positive values, but following the advice of J. de la Cruz Rodríguez (*private communication*) we remove this limitation as negative absorption can occur with sufficient stimulated emission. The differences between SNAPI and the other methods around  $\Delta\lambda = 0.02$  nm is also likely due to choices made when estimating the derivatives and limiting the control points. SNAPI also uses a different formulation for the background emissivities and opacities, as well as different parametrisations for the damping terms present in the Voigt profile. These are likely responsible for the differences in the  $|\Delta\lambda| \geq 0.15$  nm far wing to continuum region. Nevertheless, we consider that the agreement between the three implementations is very good on this challenging test.

Fig. 4.3 shows a simple validation test for the PRD method implemented in *Lightweaver*. We show the synthesis of the Ca II K line in statistical equilibrium from a snapshot of a RADYN simulation. This snapshot is taken from the F9 simulation discussed in Chap. 5, at 5 s after flare heating starts. The CRD solutions are shown in blue and dashed orange for RH and *Lightweaver* respectively. The PRD solutions both employ angle-averaged PRD and are presented in green (RH) and dashed red (*Lightweaver*). The agreement between these solutions is again very good, with a slight observable difference around  $|\Delta\lambda| = 0.012$  nm where the line in *Lightweaver's* PRD solution appears slightly wider than in RH's. This difference is present only over a very small range of wavelengths and we consider the agreement between the two implementations to be very good.

Slightly more complex examples are needed to test the time-dependent machinery present in *Lightweaver*, and neither SNAPI nor RH, with their time-independent viewpoints can be

## 4.6 Validation

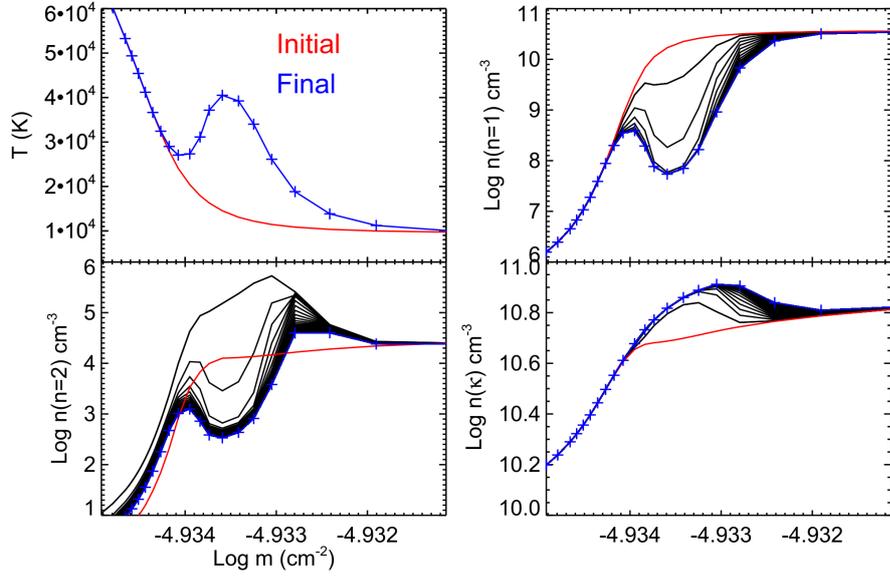


**Figure 4.3:** Comparison of *Lightweaver* and RH synthesis of Ca II K from a RADYN snapshot comparing the effects of PRD.

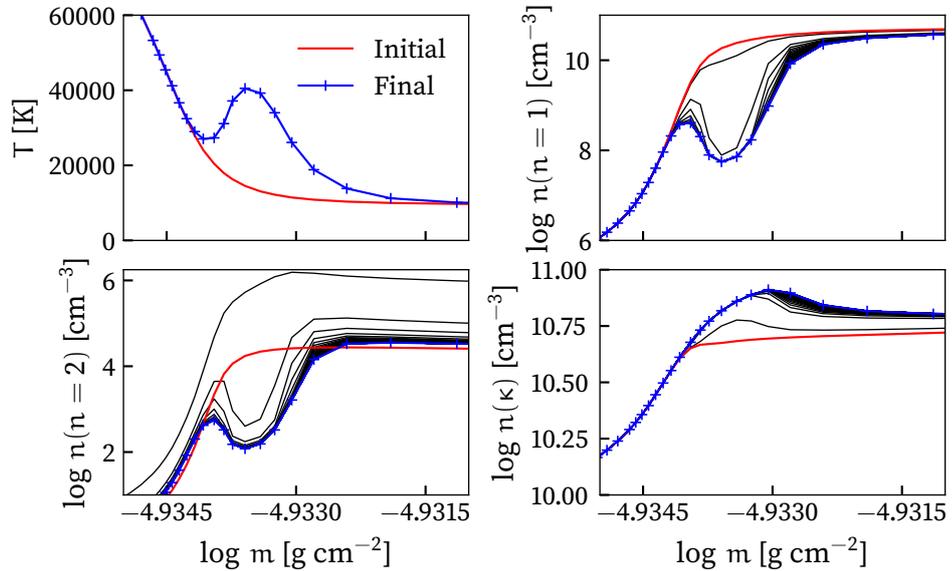
used for comparison here. We instead use the example of a perturbed FALC atmosphere presented by Judge (2017). This figure is reproduced here as Fig. 4.4. After converging to the statistical equilibrium solution in a standard FALC atmosphere, the temperature is perturbed as shown by the blue line of the top-left panel of both Fig. 4.4 and our own solution is presented in Fig. 4.5. This model uses a three level plus continuum model hydrogen atom, and the populations of the ground, first excited, and continuum states are shown in the top-right, bottom-left, and bottom-right panels of these figures respectively. The red line shows their starting values, and the blue line their final values after the simulation has been allowed to run for 500 s in the case of Fig. 4.4 and 60 s for Fig. 4.5, as we find the solution has stabilised by this point. The black lines represent the solutions for each population every 1 s. The  $x$ -axis on these plots is the atmospheric column mass, and Fig. 4.5 is prepared in cgs to allow direct comparison to Fig. 4.4. The typically temperature-dependent (i.e. due to assuming a Maxwellian electron distribution) collisional ionisation and excitation “strengths” are fixed to their values at 7000 K obtained using the method of Johnson (1972) (P. Judge, *private communication*; we note that the collisional *rates* associated with these still scale with  $\sqrt{T}$ ).

We see good overall agreement with Judge (2017), but the differences are larger than those presented in the previous figures. The method used by Judge (2017) is different to ours, and

#### 4.6 Validation



**Figure 4.4:** Fig. 4 of Judge (2017). The response of hydrogen levels ( $n = 1, 2$ ) and continuum ( $\kappa$ ) are shown for an instantaneous perturbation of the temperature of the FALC model (top-left panel). Each black line shows the population densities once every 1 s. Note that the abscissa should be labelled as  $g \text{ cm}^{-2}$ . © AAS. Reproduced with permission.



**Figure 4.5:** Validation of the time-dependent population update scheme in *Lightweaver*. Compare with Fig. 4.4. Each black line shows the population densities once every 1 s.

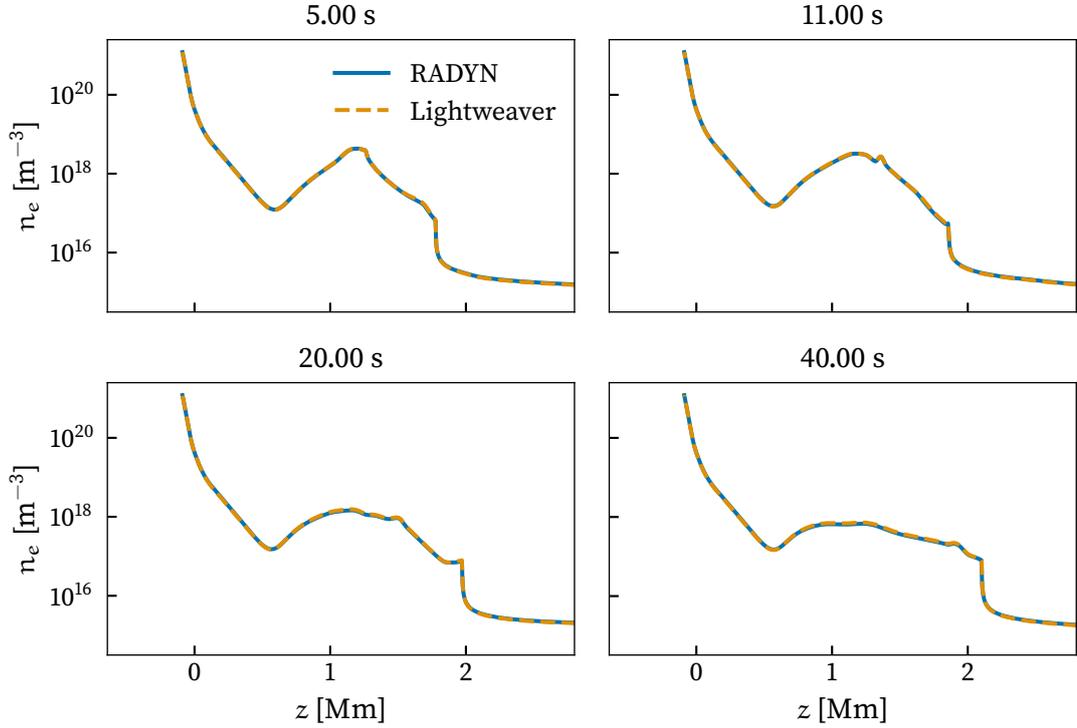
## 4.6 Validation

uses full preconditioning but with a  $\Lambda$  operator based on the escape probability formalism of Hummer & Rybicki (1982), using a one-sided escape probability. This approach is much more approximate than the full MALI treatment applied in *Lightweaver*, as it only performs the approximate formal solution at one wavelength per transition, and computes the integrals over wavelength and angle analytically. The advantage of this treatment is that it is much less computationally intensive. Thus, the use of this approximate method by Judge (2017) is likely to be the origin of the differences between our results, and we find that the level populations converge to similar final solutions, at apparently similar rates. The most substantial difference between our results is for the  $n = 2$  population: in Fig. 4.4, the points close to a column mass of  $-4.9315 \text{ g cm}^{-2}$  do not vary from their initial values, whereas in our model they change quite dramatically, with an enhancement of up to 2 dex, leaving the range of the original plot, but converge close to the expected final solution. This is likely an effect of downgoing radiation not being considered in the escape probability based formal solution of Judge (2017). The maximum enhancement in this level also peaks higher than that seen in Fig. 4.4, but rapidly drops towards the expected solution. The downgoing radiation is also likely responsible for the offset of the final populations from the initial model in the  $-4.9315 \text{ g cm}^{-2}$  region that is not present in Fig. 4.4, as the populations stay stable when the populations are advanced in time through the same process without the temperature perturbation, i.e. statistical equilibrium is maintained. This simple example illustrates that *Lightweaver's* behaviour is reasonable when applied to time-dependent problems, and the quality of its treatment will become apparent in the in-depth comparisons with RADYN presented in Chap. 5.

In Fig. 4.6 we show the electron density in the lower atmosphere at four different timesteps of a RADYN simulation where the radiative transfer (including self-consistent electron density) has been reprocessed using *Lightweaver*. The techniques used here will be described in depth in Chap. 5, and this figure is produced using the F9 model described in Sec. 5.4.1. The *Lightweaver* model presented is run with time-dependent charge conservation, maintaining self-consistency throughout the simulation, whilst loading the other thermodynamic parameters from the RADYN model at each timestep, and using these to compute the time-dependent population updates and self-consistent electron density. Advection of the atomic populations and electron density is performed using a similar technique to RADYN, implemented as per Sec. 5.3.1. The electron density in the RADYN model is shown in blue, and the *Lightweaver* model in dashed orange. The agreement between the two models is extremely good, including the fine features around  $z = 1.2 \text{ Mm}$  shown in the 11 s panel.

*Lightweaver* agrees well with the other models against which it has been tested here; it is

## 4.7 Lightspinner



**Figure 4.6:** Comparison of electron density in a RADYN simulation and time-dependent re-processing using *Lightweaver* that will be presented as the F9 model of Chap. 5.

easy to construct new validation tests for statistical equilibrium cases that can be run with many different extant radiative transfer codes, but the validation of time-dependent treatments on their own is more difficult, due to current tools often coupling these equations to hydrodynamics. Other tests have been undertaken to verify the performance of *Lightweaver*, but those presented here should be sufficient to demonstrate its capabilities.

### 4.7 *Lightspinner*

During the development of *Lightweaver*, a simpler pedagogic framework was also constructed. This framework, *Lightspinner*<sup>7</sup> (Osborne 2020), is written in pure Python and focuses on documenting the internal numerics of a simple formal solver and the MALI method using full preconditioning (under the assumption of CRD). It is accompanied by a slide deck highlighting the most important terms that need to be understood to implement

<sup>7</sup>Available on GitHub (<https://github.com/Goobley/Lightspinner>), with archival on Zenodo.

## 4.8 Discussions

these methods following Rybicki & Hummer (1992) and Uitenbroek (2001) for iteration, and Olson & Kunasz (1987) and Auer & Paletou (1994) for the short-characteristics formal solver (although only a linear formal solver is implemented in the code). This framework can be employed to help users familiarise themselves with the concepts of NLTE radiative transfer and some of the techniques present in *Lightweaver*: it presents the core concepts clearly, and naïvely, without focusing on performance, so can easily be dismantled and understood by a single person over the course of a few days.

## 4.8 Discussions

We have presented a description and validation of the *Lightweaver* radiative transfer framework and the intentions behind its design. Frameworks can substantially enhance productivity, and enable the construction of specialised tools without the need to focus on the implementation or performance of the common core of dense numerical code shared by programs of this style. The power of this will be demonstrated with the experiments presented in Chapters 5 and 6 which leverage *Lightweaver* significantly, and demonstrate how the addition of small amounts of Python can yield tools that would otherwise require in-depth modification and coupling of existing tools such as RH and RADYN.

# 5

## Radiative Transfer with Time-Dependent Populations

Modern Radiation Hydrodynamic (RHD) codes as described in Chap. 2 are highly complex, and contain many specialised features. In the following discussion we will focus primarily on RADYN, the most widely used code of its ilk, and show how using additional tools can facilitate new avenues of investigation.

### 5.1 A Brief Dissection of RADYN and a Possible Future of RHD Modelling

*This section is informed by my discussions with Prof. Mats Carlsson of the University of Oslo, experiences using RADYN, and analysis of its source code. It represents my own conclusions from the synthesis of these.*

RADYN's design closely follows its radiative transfer lineage. Its direct predecessor is the MULTI radiative transfer code (Carlsson 1986, 1992) and many commonalities remain. NLTE radiative transfer is solved on a per transition basis using an ALI method and linearisation of the resultant kinetic equilibrium equations. This method solves for the case of non-overlapping lines but includes an underlying background continuum. This linearisation approach was proven by Socas-Navarro & Trujillo Bueno (1997) to be effectively equivalent to that of preconditioning for non-overlapping transitions (i.e. MALI, Rybicki & Hummer 1991) for pure radiative transfer problems in the statistical equilibrium case. Their analysis assumed the use of a local diagonal  $\Lambda^*$  operator; RADYN, however, chooses to employ a pentadiagonal  $\Lambda^*$  operator to make optimal use of matrix bandwidth needed elsewhere in the program and obtain improved convergence as a result. It is unlikely that

### 5.1 A Brief Dissection of RADYN and a Possible Future of RHD Modelling

this change in operator significantly affects the conclusions of Socas-Navarro & Trujillo Bueno (1997), but the two methods will no longer arrive at exactly equivalent numerical formulations. RADYN also uses an efficient Feautrier formalism for its formal solver, which solves for both the up- and down-going rays simultaneously but cannot handle both overlapping lines and Doppler shifts. There are no lines with significant overlap considered in the standard model atoms used in RADYN, and continuum emissivities and opacities rarely change sufficiently over the wavelength range of a line to need to be considered in a wavelength varying sense. This is likely a fair trade-off for RADYN given the computational benefits it can bring.

The advantage of the linearisation approach applied in RADYN is the ability to directly couple other equations to the RTE and implicitly solve all of these simultaneously and self-consistently. Taking for example the kinetic equilibrium equation (2.11), RADYN's method formulates this expression such that the corrections from both the advection and population transition terms are considered simultaneously. This is achieved through the use of a Newton-Raphson method, where the Jacobian is computed based on an analytic derivation, including the aforementioned linearisation of the kinetic equilibrium equations. This same process simultaneously solves for the population updates, heat conduction, hydrodynamics of the system, and the new locations of the dynamic grid points on which the RHD equations are discretised following the method of Dorfi & Drury (1987).

A significant benefit of this implicit approach is a relaxation of the timestep constraints present in explicit approaches. This is particularly important when considering the very fine grid spacing often required by the dynamic grid which, combined with the large bulk velocities occurring in flares, can lead to extremely oppressive timestep constraints. Thermal conduction (which is typically stiff in an explicit treatment) is also computed as part of this system, using the flux limiter of Campbell (1984). Whilst this limiter will limit the local electron flux to below the free-streaming limit, the implementation is not physically accurate and non-local transport terms resulting from the solution of the Vlasov-Fokker-Planck equation may further affect this term, however any adjustment of the thermal conduction treatment is outside the scope of the work presented here.

Despite its elegance, there are several major downsides to this implicit approach. Foremost of these is the complexity engendered by the coupled design of the system, and the need to ensure that all necessary derivatives are analytically derived and correctly computed. This presents a very large barrier to entry for future developments on the platform, and is likely part of the reason why both Fokker-Planck modules integrated in RADYN have

## 5.2 Minority Species Modelling

operated externally to this core coupled system. Additionally, implicit codes, whilst having less severe timestep constraints, are typically much more costly per timestep than explicit codes. This is somewhat offset by the majority of the cost of each step residing within the formal solver, which remains similar in both cases. The dynamic grid can also become problematic due to the difficulty interpreting and manipulating the locations of the grid points. The positions of these points are determined by a metric taking into account local variations and concentrations of multiple quantities, with manually determined weightings. These difficulties are linked to a tendency to drop to extremely fine spacings in shock regions, severely limiting the possible timestep (sub-micron spacings have been observed in non-convergent simulations). Due to the fixed number of grid points used throughout the model, this reduces spatial accuracy elsewhere in the simulation. The grid weights can be manually adjusted during a run (e.g. [Kowalski et al. 2015](#)), but a reliable, fully automated process is much more desirable.

RADYN is a fantastic tool that has enabled insight into many different flare-associated phenomena, and these comments merely intend to highlight avenues for future development within the field of RHD. As the different applications of RADYN continue to evolve in complexity, with projects such as multi-strand arcade and minority species modelling (e.g. [Kerr et al. 2019a](#); [Polito et al. 2019](#); [Kerr et al. 2020](#)), the code at the core of RADYN will need to be modified by different researchers, and work facilitating this and highlighting additional factors to be considered in RHD modelling is key to the future development of this field.

As discussed previously with respect to *Lightweaver* and radiative transfer, a flexible framework designed for solving a class of problem can yield significant advances in productivity. The task of designing, constructing, and testing a framework for a problem as complex as the complete quasi-one-dimensional RHD simulation of flares is too significant to be undertaken here. Nevertheless, it may prove a powerful future development once the necessary specifications are defined. In the following we focus on reprocessing aspects of the radiative transfer of previously computed RADYN simulations and investigating important directions for future developments in RHD modelling of flares.

## 5.2 Minority Species Modelling

For flares, RADYN's primary focus is on the major spectral lines and continua of hydrogen, helium, and calcium. These typically represent the bulk of the radiative energy lost from

## 5.2 Minority Species Modelling

the chromosphere. Singly ionised magnesium has also been shown to be an important contributor to these energy losses, however the h and k lines require a treatment including PRD to avoid significantly overestimating their losses. The Ca II H and K lines are also somewhat affected by PRD, in addition to the hydrogen Lyman lines. For the Lyman lines we will discuss strategies for approximating this treatment. For Ca II H and K, it has been suggested that considering the radiative losses of these lines in CRD approximately accounts for the lack of detailed Mg II h and k treatment if all of these transitions were treated with PRD (Carlsson & Stein 2002; Kerr et al. 2019b).

Whilst the lines of these four species, H, He, Mg, and Ca are some of the strongest in the solar spectrum, and their continua mediate much of the energy leaving the chromosphere, there are other chromospheric transitions that can be used to diagnose the atmosphere. An element treated as a “minority species” is assumed to not interact significantly with the energy balance of the simulation (i.e. the thermodynamic response of the model does not change significantly if this species is subject to a complete radiative treatment). This should be true for most species with trace populations. For example, optically thick Si IV formation in a heated chromosphere has been investigated in a minority species context by Kerr et al. (2019c). The radiative transfer calculations associated with a minority species can then be performed in a “second-pass” over a previously computed RADYN simulation. The MS\_RADYN code is a modification of RADYN designed for this task; it takes the thermodynamic parameters from every timestep of a RADYN simulation, along with the non-equilibrium hydrogen populations, and solves the kinetic equilibrium equation at each timestep for a minority species. Due to the lack of atmospheric thermodynamic response to changes in the radiative output of this species, far more complex atomic models can be used, such as the 30 level model silicon atom used by Kerr et al. (2019c).

An approach similar to that of minority species modelling can be applied to testing the methods used in RADYN, the importance of certain omissions, and the feasibility of extensions. Due to the reduced complexity of solving the kinetic equilibrium equations rather than the entire RHD system, these calculations typically run significantly faster than the original simulation. In the following we will discuss the creation of a minority species tool for reprocessing RADYN simulations, built on the *Lightweaver* framework, as well as its application to investigating the importance of overlapping transitions, and discuss the difficulties of including PRD in these simulations. Building such a tool on the *Lightweaver* framework should provide researchers with a modern, simpler codebase that is easier to conceptualise and modify, allowing for investigation of effects to be included in RADYN or future RHD codes. Excluding the model atom definitions, the source code of

### 5.3 Reprocessing RADYN Simulations with the *Lightweaver* Framework

the simulations presented in this chapter totals ~1000 lines of Python, mostly following modern best practices.

This approach could also be applied to a simple investigation of the effects of a magnetic field on the outgoing radiation, by imposing an *ad hoc* magnetic field (constant or varying) on the model atmosphere and computing the emergent line profiles, in a much simpler way than modifying RADYN for this task.

### 5.3 Reprocessing RADYN Simulations with the *Lightweaver* Framework

To perform a minority species simulation, a particular file from the original simulation, `atmost.dat`, must be provided. From investigating the contents of this file we can determine the exact configuration of *Lightweaver* and the equations to be solved. When requested, this file is written to for every internal timestep of the RADYN simulation, and represents a limited subset of the less frequently written “complete” output (typically stored every 0.1 s). It contains some metadata describing the size of the simulation, then for each internal timestep, it records the current timestep, the elapsed time, the current locations of the dynamic grid, the mass density profile, the electron density profile, the temperature structure, the vertical velocity, and the current hydrogen level populations.

For the validation of *Lightweaver* and this style of simulation, we also wish to compute and compare the hydrogen populations to those computed in RADYN. Several difficulties arise due to the non-thermal collisional rates used in the kinetic equilibrium calculation for hydrogen and helium. The non-thermal collisional rates of Fang et al. (1993) are used to determine hydrogen ionisation, and require the beam energy deposition throughout the atmosphere at each timestep. For helium, if the Fokker-Planck electron beam description is used, then the rates of Arnaud & Rothenflug (1985) are used, but these require integration over the electron energy distribution. Whilst it is possible to add both of these to the `atmost.dat` file, the complete electron distribution information is very large, and we instead elect to use “Emslie” beam electron formalism (Emslie 1978), for which the energy deposition profile throughout the atmosphere is sufficient to describe the non-thermal rates. We therefore chose to slightly modify one of RADYN’s output routines, and add the beam deposition profile to the `atmost.dat` file. Our function for reading these files can handle files written both with and without this modification. In the event that this beam heating information is not saved, an approximation of it can be reconstructed via

### 5.3 Reprocessing RADYN Simulations with the *Lightweaver* Framework

interpolation from the information in RADYN's complete save file. Our testing of this show that the approximation is relatively good, but short term, or particularly narrow heating features may be lost. For this reason, all simulations presented here use the version with the beam energy deposition data.

With the above data we have sufficient information to construct the RADYN thermodynamic atmosphere at any of its internal timesteps. *Lightweaver* is then responsible for determining the radiation field and atomic level populations throughout the atmosphere using the thermodynamic properties at each timestep loaded from the RADYN output and the populations at the previous timestep. Thus, the atomic level populations are computed in statistical equilibrium for the RADYN starting atmosphere and are then evolved only by our simulation code using *Lightweaver*, and not RADYN, which simply provides the thermodynamic structure of the atmosphere at each timestep. *Lightweaver* does not make use of the mass density stored by RADYN directly, but instead maps it to hydrogen density. For this we use the default abundances in *Lightweaver*, based on [Asplund et al. \(2009\)](#). These differ to those used in RADYN, but not significantly for any of the species discussed here.

Ignoring the advection term it is then simple to produce a minority species tool using this approach. In many situations, the advection term has a small effect, and can safely be ignored ([Kašparová et al. 2003](#); [Nejezchleba 1998](#)). The method for advancing the atomic populations in time employed in RADYN is formulated on the dynamic grid, and this is not the case in *Lightweaver*, which assumes that the grid is static (although this limitation can be worked around). We can simply use a fixed denser spatial sampling of the atmosphere to account for the motion of features such as the transition region over the course of the simulation. This model then interpolates the thermodynamic properties and NLTE hydrogen populations for the starting atmosphere onto our stratification and computes the statistical equilibrium solution for the minority species in question. For each subsequent timestep these properties are interpolated from the new RADYN grid to the static grid, and the minority species populations can be advanced in time using the process described in [Sec. 2.4.11](#).

To reduce the number of grid points needed for a static stratification one could instead use a fixed column mass stratification. The transition region moves very little in terms of column mass during the simulation, however it then becomes necessary to interpolate the populations from one column mass stratification to the next; a process which can introduce significant error if not undertaken with care. This error can be reduced by renormalising each species' total number density throughout the atmosphere from the

### 5.3 Reprocessing RADYN Simulations with the Lightweaver Framework

mass density and abundances, and this helps to avoid errors growing around regions of high gradient, such as the transition region.

All of these approaches were trialled with different RADYN simulations, and it was clear that while the general evolution of the radiative output was similar, differences remained in the synthesised line profiles, especially those of Ca II. To solve the minority species problem properly in a manner fully compatible with RADYN it is necessary to include the advection terms, and so an advection treatment was added. In the following we will describe two different approaches to handling the advective terms, drawing on our previous discussions of hydrodynamics, and thus solving the complete system of kinetic equilibrium equations.

#### 5.3.1 Advection

In Sec. 5.1 we discussed the coupled manner in which RADYN solves the RHD equations. For flexibility we wish to decouple the advection terms from the radiative effects. Our initial approach was to use an explicit method on the dense, fixed grid discussed above. This method draws the simple explicit finite-volume approaches discussed in Chap. 2, and thus uses a fifth-order WENO-NM scheme for reconstruction and the simple local Lax-Friedrichs flux for estimating the flow of the atomic populations through the simulation. This numerical scheme performs extremely well in textbook advection and hydrodynamic tests, however it was found to be ill-suited to the long internal timesteps that RADYN's implicit method chooses. These timesteps are longer than those permitted by the CFL condition, requiring the explicit method to perform multiple substeps for each of RADYN's timesteps. This alone can introduce cumulative error, and is further compounded by the use of the dense static grid. As this grid needs to conform to the requirements of the simulation over its entire evolution, rather than just its instantaneous requirements (as is the case for RADYN's grid) there are often dense clusters of points in regions where they are not currently required to resolve the atmospheric structure. These regions may have high plasma flow velocities, further increasing the timestep restrictions on the explicit method. In some cases several thousand applications of the explicit scheme were required to match one of RADYN's internal timesteps. With such different timestep restrictions imposed on the two components of this simulation, it is difficult to correctly treat both processes in a coupled fashion.

Using this explicit approach for advection, a fair agreement with RADYN was found, but performance was dramatically worsened. A significant proportion of runtime had

### 5.3 Reprocessing RADYN Simulations with the Lightweaver Framework

to be spent on the advective terms as the energy deposition in simulations increased, which increased the flow speeds and moved the transition region further from its starting altitude. Due to RADYN's longer timesteps, during more explosive evaporative phases, our advection scheme was occasionally able to carry chromospheric material through the majority of the transition region. Whilst the populations quickly evolved in their new environment this process is sufficiently slow to introduce unexpected variability in the line profiles. A higher order splitting scheme might have helped to mitigate this but would add further computational cost to an already expensive process (due to the sheer number of advection steps needed). This effect can also be mitigated by limiting the maximum timestep that can be taken by RADYN during the creation of the baseline simulation. However, limiting RADYN's performance is a poor solution to this problem; the explicit advection method described here could serve well at the core of a code designed around it, in the way that FLARIX and HYDRAD both use similar explicit advection schemes, but many difficulties are created by the difference in ideologies between this and RADYN's implicit dynamic grid.

To improve the agreement with RADYN, reduce the number of interpolations, and leverage rather than fight its adaptive grid we instead decided to apply its technique for advection, but keep it distinct from the radiative transfer. Thus, at the start of each timestep we advect the populations from the previous grid to the new grid locations, and then advance the populations in time based on the NLTE rates. This method employs a variant of the second order spatially accurate method of [van Leer \(1979\)](#) to reconstruct the values on either side of the interface, of which the upwind value is chosen. This is set within a time-centering scheme similar to that described by [Dorfi \(1997\)](#). Under this scheme the advection stencil depends on the values of five cells, two on either side of the current one. Thus the Jacobian matrix for this equation is pentadiagonal (and block pentadiagonal for the complete system of ODEs). For simplicity and ease of implementation we chose to not directly code the expressions for these derivatives that occur in the Jacobian matrix, but instead compute them by finite differences. There is an elegant technique that can be used to optimise this process when an equation's region of dependence is known. This approach is known as Coloured Finite Difference ([Curtis et al. 1974](#)); so long as only one point in the domain affecting each output point is perturbed the gradient can be trivially attributed to the correct point. In essence, we need only perform five additional evaluations of the advection residual, perturbing one point in every five, to fill the Jacobian, rather than the traditional approach where each point would be perturbed in turn. The solution to this system is then computed by a Newton-Raphson iteration with Armijo line search ([Armijo 1966](#)) to accelerate convergence.

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines

This technique has proven far more suited for reprocessing these simulations than the previous explicit scheme. Only one application of the method (typically requiring at most five Newton-Raphson iterations) is needed for each timestep and the computational cost is typically  $< 0.1\%$  of the total CPU time. This approach remains harder to debug and modify due to its implicit nature, although this is simplified thanks to the finite difference method for computing the Jacobian which does not need to be adjusted if the equations are modified, so long as the stencil remains the same.

The agreement in the synthesised lines was improved by the implementation of advection in our simulations, but substantial differences still remained in the spectral lines of Ca II. In the following, we investigate the effects of different radiative treatments, and discuss possible implications on RHD modelling.

This final code used in these investigations is available on GitHub<sup>1</sup> with archival on Zenodo (Osborne 2021b). The code for the previous advection method on the dense fixed grid is available on the `Advection` branch of the same repository. This is an early development branch of the code, as this method was not retained.

### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines

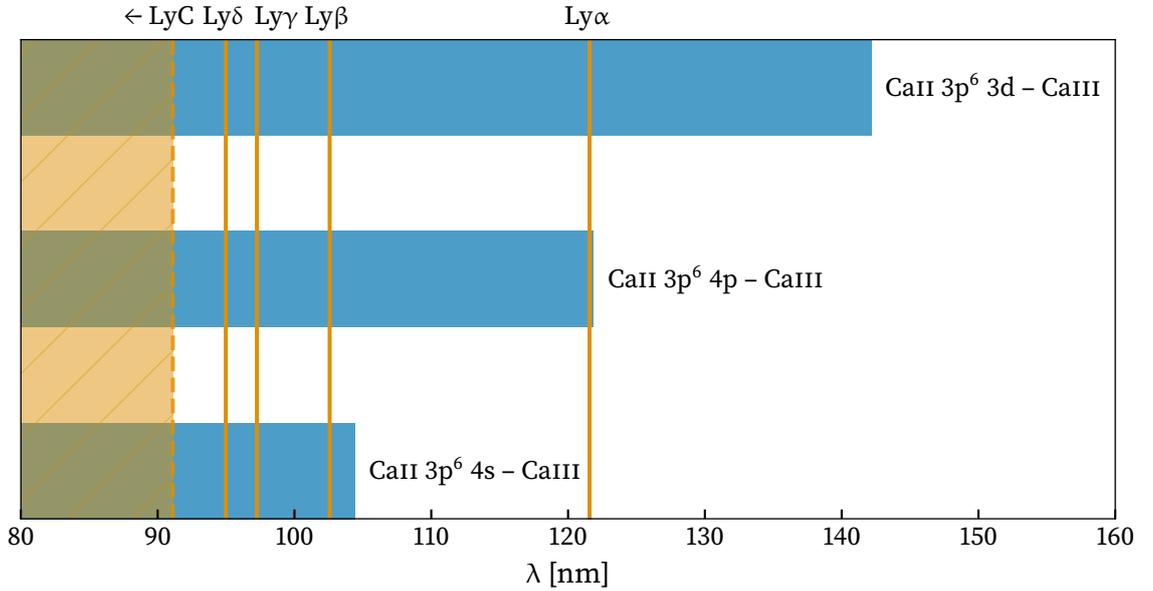
*The content of this section is based on the work presented in Osborne et al. (2021).*

The methods discussed above can be used to reprocess the radiative transfer aspects of RADYN simulations and thus investigate the effects of the Lyman lines on Ca II photoionisation in these models, as this was theorised to be the origin of possible differences in the Ca II line profiles. In RADYN, the photoionisation of Ca II by the hydrogen Lyman continuum is considered, but the effects of the Lyman lines are not. The hydrogen Lyman lines are strongly enhanced in flares, and a two dex enhancement relative to quiet sun intensity was found by Rubio Da Costa et al. (2009) using observations from the Transition Region and Coronal Explorer (TRACE). RHD modelling using RADYN has also suggested that enhancements at least this large are to be expected in this line (Brown et al. 2018; Hong et al. 2019). Fig. 5.1 shows the overlap between the hydrogen Lyman transitions and the Ca II continua present in our model. Radiation from Ly $\alpha$  photoionises Ca II from all levels other than the ground state. All of the other Lyman transitions present can

---

<sup>1</sup><https://github.com/Goobley/MsLightweaver>

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines



**Figure 5.1:** The overlap between the hydrogen Lyman lines and continuum with the Ca II continua present in the model atom used here. Both the Ca II  $3p^6 4p$  and  $3p^6 3d$  levels contain two sub-levels with indistinguishably different continuum edges.

photoionise Ca II to Ca III from all levels present in this model. Additionally, in flaring conditions, the higher lines of the Lyman sequence are significantly Stark broadened and create a quasi-continuum between Ly $\delta$  and the Lyman continuum (De Feiter & Švestka 1975), which will further enhance the photoionisation of Ca II from what is considered in our model. These highly enhanced transitions therefore provide an important mechanism for the photoionisation of Ca II to Ca III and in the following we will investigate the effect this has on both the emergent line profiles and the radiative losses of several RHD models. These effects have long been considered in prominence modelling, starting with the work of Ishizawa (1971), however we are not aware of any previous detailed investigation of these effects in RHD modelling of flares, where the Lyman transitions become so significantly enhanced.

The hydrogen (five bound levels with H II continuum and ten lines), helium, and calcium (five bound Ca II levels with Ca III continuum and five lines) model atoms used in the simulations presented here are the same as those used in RADYN. All other model atoms used in the LTE background opacities are taken from the *Lightweaver* standard library (and have in turn been converted from RH distribution (Uitenbroek 2001)). Currently only the hydrogen and calcium populations are fully treated in NLTE. In the RADYN models the tabulated LTE background emissivities and opacities of Gustafsson (1973) are used.

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines

To determine the importance of the photoionisation by the Lyman lines, each RADYN simulation is reprocessed twice, once including the effects of the Lyman lines on Ca II (henceforth Lyman inclusive (LI)), and once without these effects (Lyman exclusive (LE)). This is achieved through the use of two different hydrogen model atoms, where the one for the LE treatment excludes the Lyman lines and uses fixed hydrogen populations from RADYN. In both cases the effects of the Lyman continuum are included, as this is commonly considered in RHD codes, and allows for direct comparison of the LE simulation against RADYN. The simulations are performed using the CRD formalism for consistency with RADYN, although the models used for the Lyman lines contain an approximation to PRD by removing radiative broadening and reducing van der Waals broadening. This makes the line profile closer to a Doppler profile, by reducing the importance of the Lorentzian wings. This is one of the two more common approaches to approximating PRD effects in these transitions, the other being to truncate the line quadrature around ten Doppler widths from the line core. Both of these are empirical but we favour the former in flare simulations as the narrow grid used in the latter can easily “lose” opacity in the moderate to high Doppler shifts that occur in flare models (M. Carlsson 2021, *private communication*). PRD effects are also present in the Ca II lines — primarily the resonance lines, although cross-redistribution effects also affect the infra-red triplet. The effects on these are typically less significant than those on the hydrogen Lyman lines, especially due to the high chromospheric densities that occur during flares. We therefore do not add any approximate PRD treatments to the Ca II lines. These model atoms are identical between RADYN and *Lightweaver*, and the two tools differ only in their description of LTE background opacities, which in the case of the *Lightweaver* simulations includes helium. Comparison of these simulations should therefore enable the study of these photoionisation effects.

Our *Lightweaver*-based tool uses the first-order splitting technique discussed in Sec. 2.5.5 to temporally couple the two subproblems of the kinetic equilibrium equations. The complete kinetic equilibrium equations are split into radiative and advective operators that are applied sequentially. A Strang-splitting approach was trialled but provided no noticeable difference in the solution over this first-order technique, whilst being more computationally costly than the simpler approach. After solving the statistical equilibrium problem to determine an initial solution for the atomic populations in the given starting atmosphere, the populations are then advected, and the thermodynamic atmospheric parameters are updated using the RADYN data from the next timestep. This update process includes the calculation of the LTE populations of all species considered and computing the resultant background opacities. The atomic level populations are then advanced in

time using the implicit approach described in Sec. 2.4.11. In the LI case, the simulations are performed with charge conservation as described in Sec. 2.4.13, and the electron density is advected with the atomic populations. No difference in the radiative output was found between our charge conservation implementation, and that found by using the RADYN values at each step, showing that this method is self-consistent. The results of this were shown during *Lightweaver* validation in Fig. 4.6 of Sec. 4.6.

### 5.4.1 The RADYN simulations

For this investigation two RADYN simulations were used. Their parameters were chosen to serve as typical simulations, and were based on those used by Kerr et al. (2019a,b). The same starting atmosphere derived from VAL3C (Vernazza et al. 1981) as used in the F-CHROMA grid of simulations<sup>2</sup> was used. As previously discussed, the “Emslie” beam formalism was chosen over the Fokker-Planck for ease of reconstructing any non-thermal rates that might be needed. The spectral index<sup>3</sup> used for the power-law distribution of electron energies was  $\delta = 5$ , with a low-energy cut-off of 20 keV. The two simulations differed only in energy deposition, which was a constant flux for 10 s of  $1 \times 10^6$  or  $1 \times 10^7 \text{ J m}^{-2} \text{ s}^{-1}$ <sup>4</sup>. These simulations will be referred to as F9 and F10 respectively.

All of these parameters fall within the range of those used in the F-CHROMA grid, other than the beam energy flux which is lower than the peak fluxes used in the grid. This is primarily due to the triangular 20 s duration heating profile used in the grid, which is less demanding on the simulation than the constant deposition used in our simulations.

Our choice of parameters are also supported by observational data. The low-energy cut-off is in accord with the findings of Sui et al. (2007) whose analysis of 33 early impulsive flares using the Ramaty High Energy Solar Spectroscopy Imager (RHESSI) found a range of 10-50 keV whilst accounting for X-ray albedo and under the assumption of a cold collisional thick target model. Another study of 53 flares using RHESSI by Saint-Hilaire et al. (2008) found that the photon spectral index  $\gamma$  was distributed between 2 and 5, peaking between 3 and 3.5. Given this, our choice of spectral index is reasonable and lies well inside this distribution as the spectral index of the electron beam is related to the photon spectral index by  $\delta = \gamma + 1$ .

---

<sup>2</sup><https://star.pst.qub.ac.uk/wiki/public/solarmodels/start.html>

<sup>3</sup>The distribution of electrons with energy  $E$  in a power-law distribution is proportional to  $E^{-\delta}$  and is controlled by the spectral index  $\delta$ .

<sup>4</sup>In the cgs units that are commonly adopted for RADYN simulations these represent  $1 \times 10^9$  and  $1 \times 10^{10} \text{ erg cm}^{-2} \text{ s}^{-1}$ .

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines

For both of these simulations RADYN's additional coronal X-ray and extreme ultraviolet (XEUV) irradiation terms were disabled due to discrepancies that were found between these and their re-implementation in *Lightweaver*. Thus, no coronal XEUV irradiation terms were used in either RADYN or *Lightweaver*, as this was necessary to ensure agreement in the LE case, allowing the differences in the LI case to be assessed. Additionally, the maximum timestep allowed in RADYN was limited to 0.01 s. This was primarily due to experiments with the explicit advection scheme, but likely also reduced the possible loss of accuracy due to the operator splitting scheme used in our final reprocessed simulations.

##### 5.4.2 Line Profiles

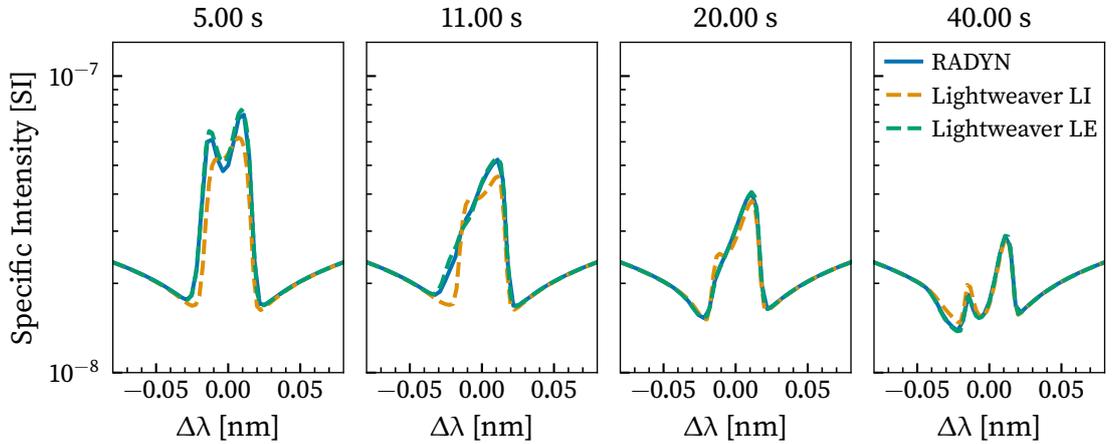
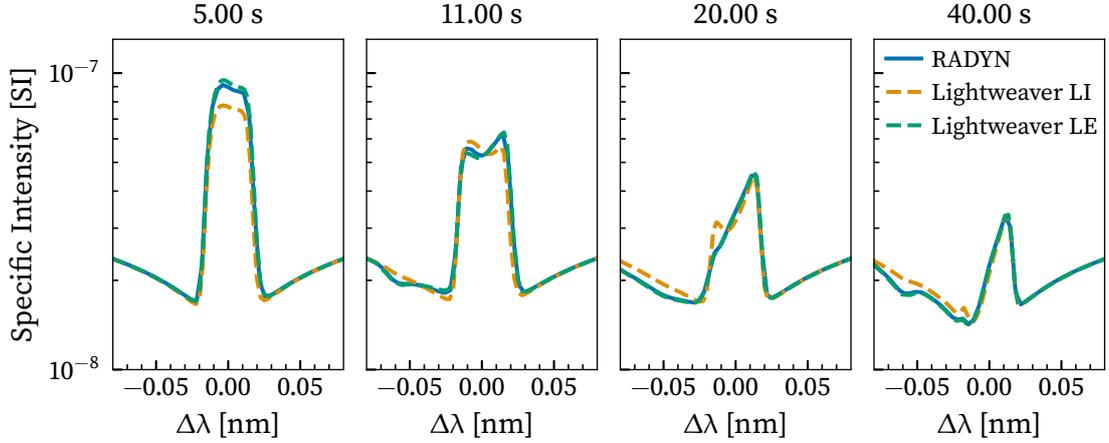


Figure 5.2: Comparison of the LI, LE, and RADYN Ca II 854.2 nm line profiles during the F9 simulation.

The Ca II 854.2 nm line profile from 5, 11, 20, and 40 s after the onset of heating in the F9 and F10 simulations is shown in Figs. 5.2 and 5.3 respectively. These figures show the comparison of the LI and LE treatments, as well as the reference line profiles computed by RADYN. Despite the different formal solvers and numerical techniques, the agreement between the RADYN and LE profiles is extremely good, differing by a few per cent at most. There are substantial differences between the LI and LE models: in the F9 simulation the LI line profile is narrower, consistently double-peaked, and less intense than the LE profile, which is much more variable, including becoming singly-peaked after the heating ends ( $t = 11$  and  $20$  s), before returning to a double-peaked shape at later times in the simulation.

The differences in the outgoing radiation in the F10 case (Fig. 5.3) are still significant,

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines



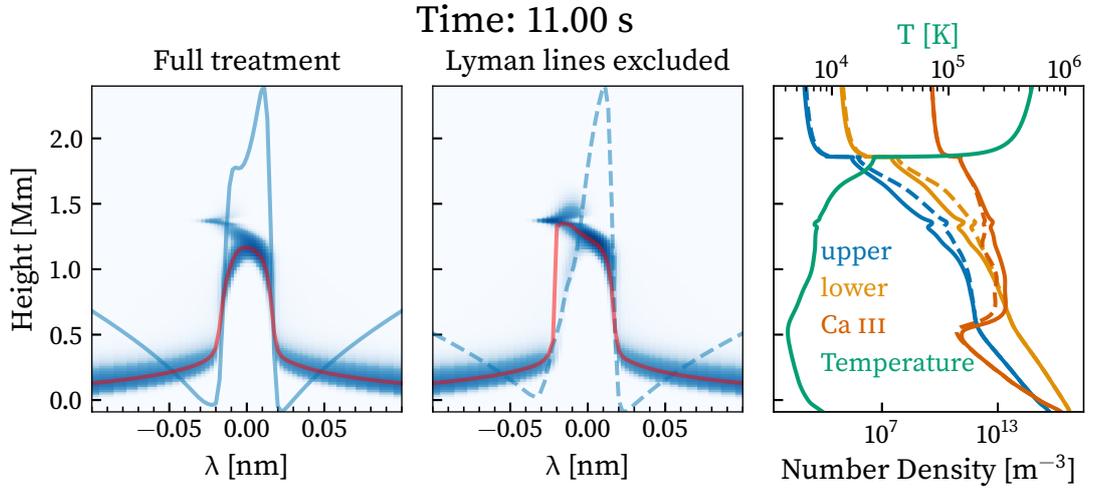
**Figure 5.3:** Comparison of the LI, LE, and RADYN Ca II 854.2 nm line profiles during the F10 simulation.

but perhaps a little less dramatic than the F9 case. The LI peak intensity is significantly reduced relative to the LE line profile in the  $t = 5$  s plot, and post-heating ( $t = 11$  s) the asymmetry of the double-peaked profiles is reversed between the LI and LE treatments. At  $t = 20$  s the situation is similar to the F9 simulation: the LE treatment produces a singly-peaked line profile, whilst the LI treatment produces an asymmetric double-peaked profile (due to the appearance of a secondary peak on the violet wing). At later times there is a significant dip in the far violet wing for the LE treatment that is not present in the LI treatment. This dip varies slowly in position and depth over the evolution of the cooling phase.

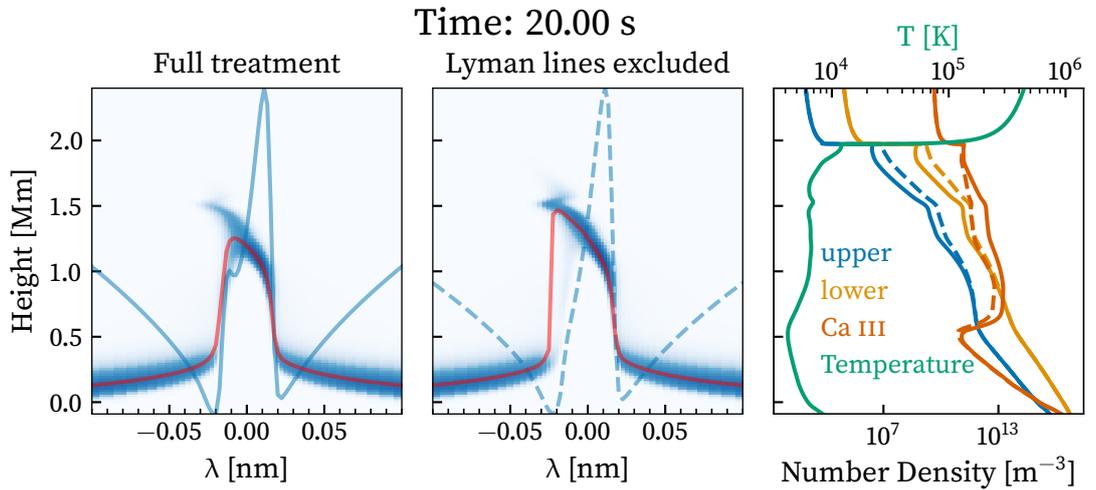
We can investigate the formation of these line profiles by looking at the contribution functions, and the atomic level populations associated with these spectral lines. From Figs. 5.2 and 5.3, we consider that the difference between the two treatments is most significant at  $t = 11$  and 20 s. The contribution functions and associated atomic level populations for these two times in both of the simulations are plotted in Figs. 5.4-5.7.

The level populations for both treatments are plotted in the right-hand panel of these figures. The LI and LE treatments produce significantly different populations for the upper and lower levels of the Ca II 854.2 nm transition. The left-hand and centre panels present the contribution functions for the two different treatments with the  $\tau_v = 1$  line overlaid in red. The  $\tau_v = 1$  line is reliably extended to higher altitudes in the line core of the LE treatment, suggesting that this core is formed in a different location between the two treatments. There is a significant feature contributing to the violet wing of the LE models around the 1.5 Mm region. In this same region both the upper and lower level populations

5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines

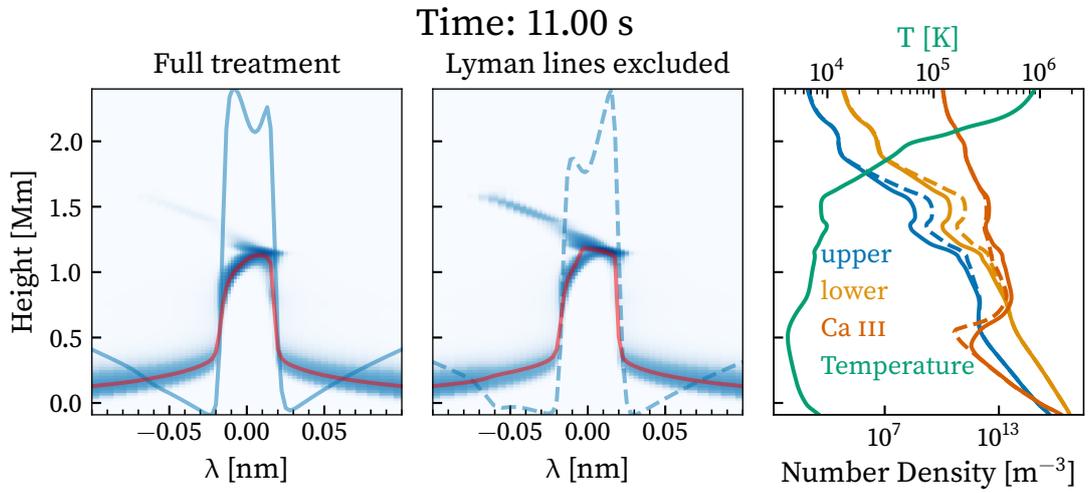


**Figure 5.4:** Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F9 simulation at  $t = 11$  s. The two left-hand panels show the contribution function, overlaid with the line profile in blue and the  $\tau_\nu = 1$  line in red. The right-hand panel shows the structure of the atmosphere, with populations from the LI treatment shown with solid lines and the LE treatment with dashed lines.

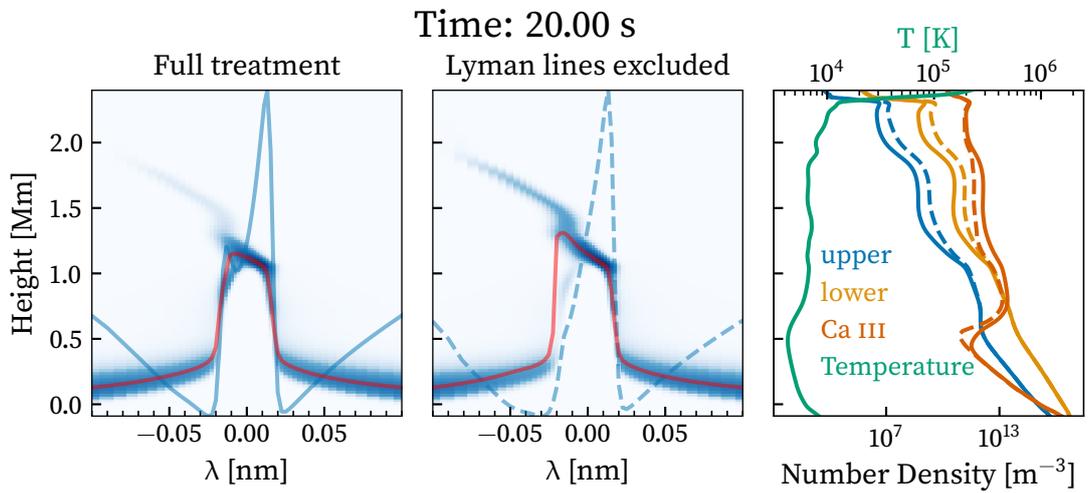


**Figure 5.5:** Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F9 simulation at  $t = 20$  s. The panels present the same information as Fig. 5.4.

5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines



**Figure 5.6:** Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F10 simulation at  $t = 11$  s. The panels present the same information as Fig. 5.4.



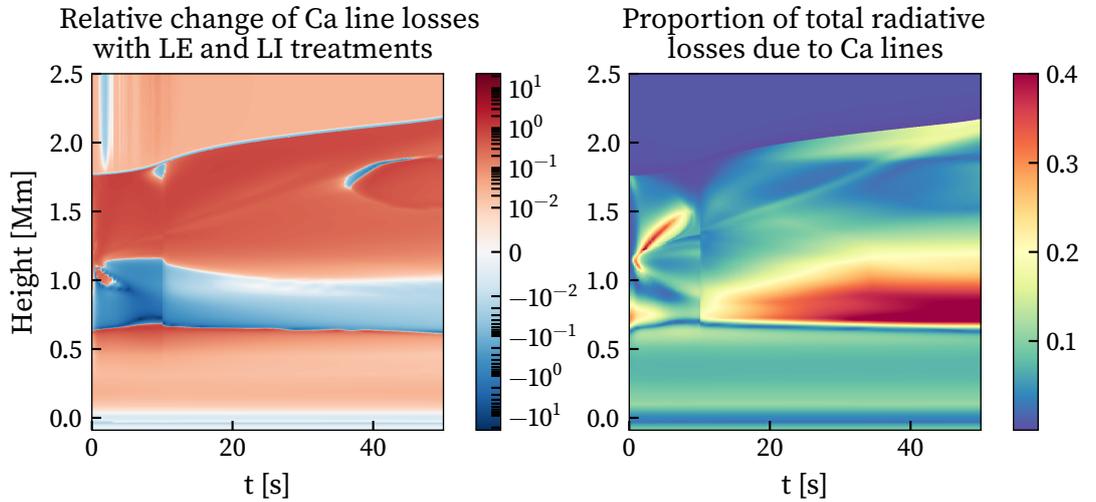
**Figure 5.7:** Ca II 854.2 nm contribution functions and level populations for the two calcium treatments in the F10 simulation at  $t = 20$  s. The panels present the same information as Fig. 5.4.

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines

for the Ca II 854.2 nm transition are significantly enhanced over their LI values, creating an increase in opacity, whilst keeping the source function approximately constant, leading to the attenuation of radiation from deeper in the atmosphere. Under the LI treatment we instead see that there is substantially more Ca III in this region, likely due to Ca II to Ca III photoionisation from Lyman lines formed in the upper chromosphere and transition region. In the F10 simulation there is also a significant difference between the Ca III populations computed in the two treatments around the temperature minimum region; this is much deeper than any of the Ca II spectral line cores typically form, but is likely due to the same photoionisation process and the effects of the much larger Lyman line intensity that occurs in the F10 model.

For both the F9 and F10 simulations, the effect of the Lyman lines on the calcium level populations (enhanced Ca III, reduced Ca II) in the upper chromosphere is similar to that found by [Gouttebroze & Heinzel \(2002\)](#) in prominence models, although the magnitude of the effect is larger here, likely due to the strong enhancement of the Lyman lines in these flare models.

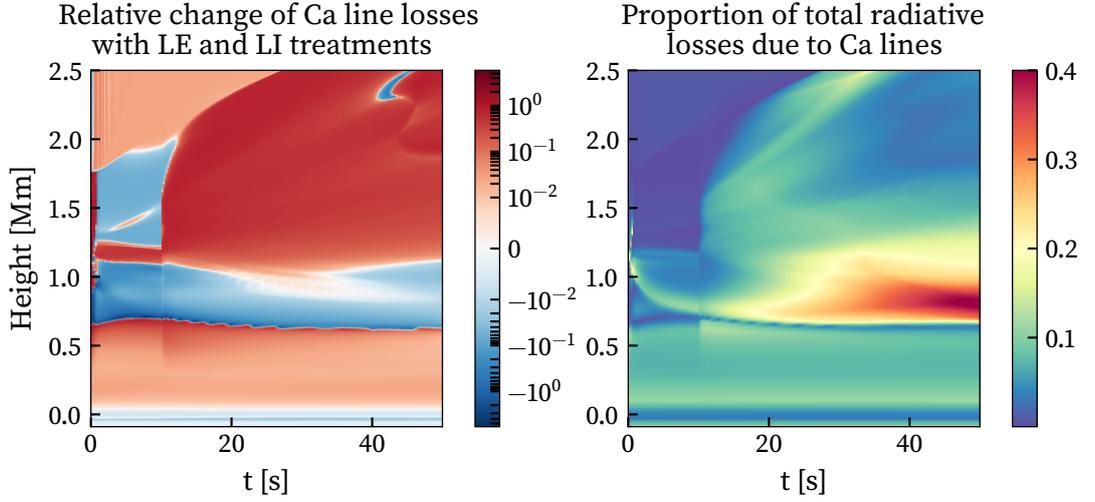
##### 5.4.3 Radiative Losses



**Figure 5.8:** Time evolution of the calcium losses in the F9 simulation. The left-hand panel shows the absolute relative change in losses due to the different calcium treatments, and the right-hand panel shows the proportion of the total radiative losses due to the calcium lines.

Whilst the term “radiative losses” is often used to describe the effects of radiation on the energy balance of the plasma, these effects are not uniquely negative. In many regions

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines



**Figure 5.9:** Time evolution of the calcium losses in the F10 simulation. The left-hand panel shows the absolute relative change in losses due to the different calcium treatments, and the right-hand panel shows the proportion of the total radiative losses due to the calcium lines.

of the plasma, absorption leads to a net gain in plasma energy, especially considering each transition on an independent basis. It is therefore difficult to immediately assess the effect that a change in radiative loss has on the system. For this we adopt an absolute relative difference metric computed following

$$\frac{\sum_{i \in \mathcal{C}} |\text{loss}_{i, \text{LE}}| - \sum_{i \in \mathcal{C}} |\text{loss}_{i, \text{LI}}|}{\sum_{i \in \mathcal{C}} |\text{loss}_{i, \text{LE}}|} \quad (5.1)$$

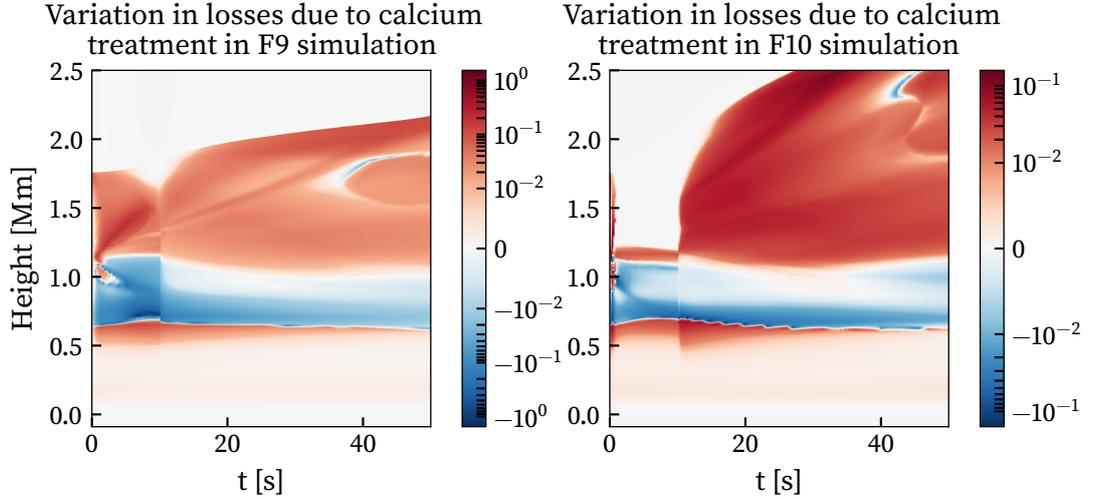
where  $\mathcal{C}$  is the set of calcium lines used on our model atom, and  $\text{loss}_i$  is the volumetric radiative loss of transition  $i$ . This metric therefore quantifies the effect on the total energy being redistributed throughout the simulation (or leaving the simulation) from the different treatments of the Ca II lines. To assess the importance of this it is also essential to know the impact of the Ca II lines on the total energy balance of the simulation. Thus we also compute

$$\frac{\sum_{i \in \mathcal{C}} |\text{loss}_{i, \text{RADYN}}|}{\sum_j |\text{loss}_{j, \text{RADYN}}|}, \quad (5.2)$$

which describes the proportion of all radiative losses considered with a detailed NLTE treatment in RADYN (hydrogen, helium, and calcium, lines and continua) due to the calcium lines. These two metrics are plotted in the left- and right-hand panels respectively of Figs. 5.8 and 5.9 for the F9 and F10 simulations respectively.

Looking first at the F9 simulation shown in Fig. 5.8, there is a significant difference between

#### 5.4 Case Study: Ca II Photoionisation by the Hydrogen Lyman Lines



**Figure 5.10:** The absolute relative change of the radiative losses considered here due to the two different calcium treatments in each of the simulations (i.e. the product of the two panels in each of Figs. 5.8 and 5.9).

the calcium losses in the two treatments during heating (0–10 s), with relative differences primarily in excess of 50 % and peaking as high as 1000 % in the 0.6–1.75 Mm region. This is not surprising given the difference in the outgoing profiles shown in Fig. 5.2. This difference remains important (typically 10–70 %) above an altitude of 1 Mm throughout the entire simulation. The right-hand panel shows that these losses represent the largest proportion of the total radiative losses in the 0.6–1 Mm region (after heating has ended). This is also the region in which the radiative losses between the two treatments agree the best, with differences below 10 % over most of this region. Nevertheless, in the region above 1 Mm where there is significant disagreement between the two treatments, the calcium losses typically represent in excess of 8 % of the total.

The effects are quite similar in the F10 simulation. During energy deposition the region where the difference between the two treatments is largest is a smaller band, which is centred on  $\sim 1.2$  Mm. The variation here is 30–50 %. This location is in agreement with the line-core formation region shown in the contribution functions. This region then expands significantly due to evaporation as the beam heating ends. Similarly to the F9 simulation, the region where the calcium lines represent the largest proportion of the total radiative losses is between 0.6 and 1 Mm at later times in the simulation, and this is once again the region where the two treatments best agree (with a similar discrepancy to the F9 case). In the region above this, where the difference between the two treatments is less but still significant, the calcium line losses typically represent 5–20 % of the total

### 5.5 Case Study: Is Full Time-Dependence Necessary?

radiative losses.

For both of the simulations presented here there appears to be a similar difference between the LI and LE calcium treatments. We can obtain a clearer estimate of the effects of the calcium treatment on the total radiative losses by plotting the product of the left- and right-hand panels of Figs. 5.8 and 5.9. These are plotted in Fig. 5.10 and confirm our conclusions. In both cases, throughout the upper chromosphere there is a variation in total radiative losses of up to 15 %. Time-averaging this variation suggests that the average difference is larger in the F9 simulation than the F10 simulation, although these are both of the same order of magnitude. This difference is likely due to a larger fraction of the total radiative losses from this region being mediated by the hydrogen transitions in the more energetic simulation, as a larger proportion of the calcium populations are ionised into the Ca III state. From these simulations, we can suggest that it is likely that using a self-consistent LI treatment (where the changes in radiative losses directly affect the hydrodynamic evolution) could produce a change in energy balance in the chromosphere of 10–15 %. This is sufficiently large to noticeably modify the atmospheric evolution, which would further affect the outgoing calcium line profiles and formation heights, but also other chromospheric lines, such as the hydrogen Balmer series.

### 5.5 Case Study: Is Full Time-Dependence Necessary?

*The work in this section was undertaken in collaboration with P. Heinzel and J. Kašparová, in parallel with the study of Ca II photoionisation, as a result of the International Space Science Institute (ISSI) meeting: “Interrogating Field-Aligned Solar Flare Models: Comparing, Contrasting and Improving” led by G.S. Kerr and V. Polito.*

It has been repeatedly shown that a time-dependent treatment of the hydrogen populations is necessary in RHD simulations. This was first investigated in wave heated chromospheric simulations by Carlsson & Stein (2002), who found a settling time for hydrogen to return to equilibrium ionisation of the order of thousands of seconds. Brown et al. (2018) also compared the hydrogen line profiles, computed with RADYN’s approximate PRD and time-dependence, and also with statistical equilibrium using the full PRD treatment of RH (i.e. treating each atmospheric subsequent snapshot in statistical equilibrium). They found that the time-dependence had a more significant effect than PRD on the Lyman lines, and could not be ignored. More recently Kerr et al. (2019b) have investigated the effects of a time-dependent treatment of the level populations of Mg II, and found that

### 5.5 Case Study: Is Full Time-Dependence Necessary?

whilst there were differences between the statistical equilibrium and time-dependent treatments, the statistical equilibrium treatment was sufficient for most stages of the flare models investigated. They conclude that for this line priority should be given to the PRD treatment. Similarly, [Leenaarts et al. \(2012a\)](#) found that for a correct synthesis of H $\alpha$  from three-dimensional RMHD models, only an equation of state taking into account the non-equilibrium ionisation and its effects on the electron density is necessary, and under this condition a full three-dimensional treatment is more important than complete time-dependence. [Wedemeyer-Böhm & Carlsson \(2011\)](#) investigated the non-equilibrium ionisation of calcium in quiet RADYN simulations and found that whilst the statistical equilibrium and time-dependent treatments gave different ionisation fractions, the difference in the synthesised line profiles was hardly discernable.

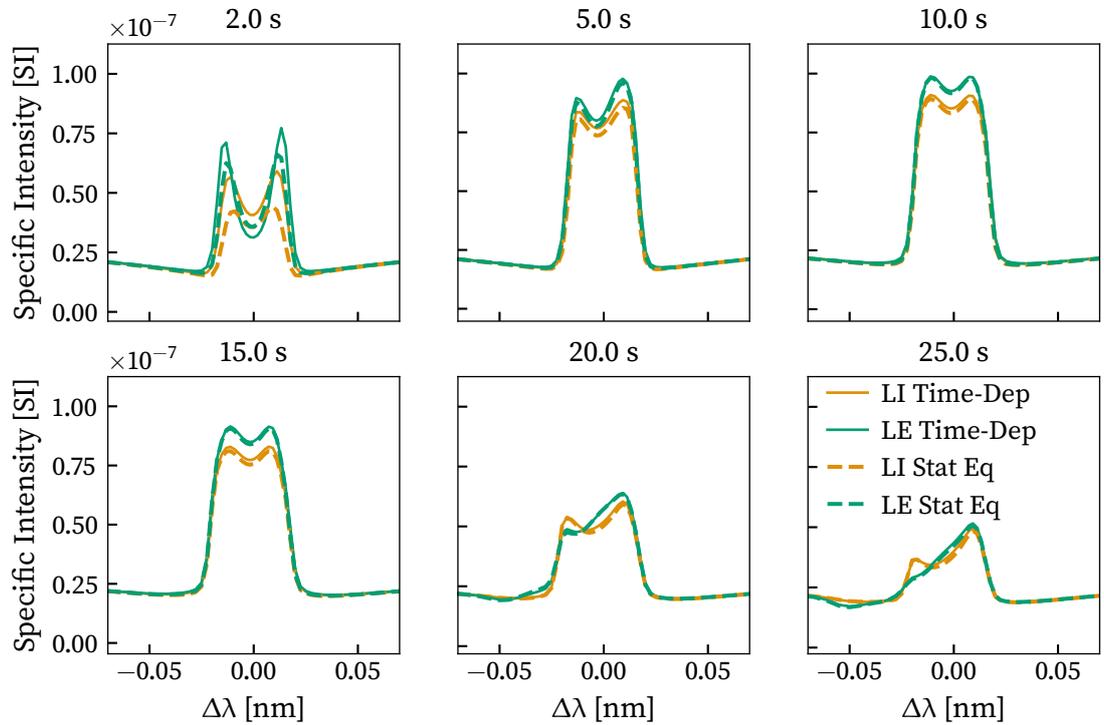
It is therefore important to assess the importance of a time-dependent treatment of Ca II in solar flare models. In the following, we shall briefly summarise our investigation of the importance of treating Ca II with time-dependence in the RHD model used for the comparison of the RADYN and FLARIX codes presented in [Kašparová et al. \(2019\)](#). This simulation utilises a set of features common to both codes, considering only hydrogen and calcium in detailed radiative transfer, and using an analytic “Emslie” beam formalism. The beam flux is modulated to have a symmetric triangular time profile, lasting 20 s and peaking at 10 s depositing a total of  $1 \times 10^8 \text{ J m}^{-2}$ <sup>5</sup>. A spectral index of 3 and low-energy cut-off of 20 keV was also chosen, and similarly to the models shown previously the initial atmosphere was based on the VAL3C model of [Vernazza et al. \(1981\)](#).

In Fig. 5.11 we present a comparison of both the LI and LE treatments of Ca II 854.2 nm line at different times in this simulation, having applied both time-dependent and statistical equilibrium treatments to the calcium populations. In all cases the hydrogen populations were loaded from a fully time-dependent run and are used in *Lightweaver’s* “detailed static” mode to provide the correct radiation field. The time-dependent treatments are considered as expected reference solutions shown with solid lines, whilst the statistical equilibrium treatments are shown with dashed lines. Comparing the time-dependent LI (orange) and LE (green) line profiles, we see a very similar picture to that presented by our previous simulations, with the peak intensity during heating being higher in the LE case and the LI model remaining more dramatically double-peaked at later times. Instead comparing the time-dependent and statistical equilibrium line profiles, we find very good agreement for all but the  $t = 2$  s snapshot, and the LI case of the  $t = 5$  s snapshot. For the latter of these, the line shape is correct, but the intensity of the statistical equilibrium

---

<sup>5</sup> $1 \times 10^{11} \text{ erg cm}^{-2}$

### 5.5 Case Study: Is Full Time-Dependence Necessary?



**Figure 5.11:** Comparison of time-dependent and statistical equilibrium treatments for the Ca II 854.2 nm line for LI and LE treatments in the simulation described in Section 5.5. The time-dependent treatments are shown with solid lines, and the statistical equilibrium treatments with dashed lines. Orange is used for the LI treatment, and green for the LE treatment.

case is a few per cent lower than that of the full time-dependent treatment. At  $t = 2$  s the difference between the two different LI treatments is greater than the difference between the LE treatments, but for both the LI and LE cases the statistical equilibrium lines are narrower and less pronounced than their time-dependent counterparts.

For this particular simulation, after the first few seconds of heating, the statistical equilibrium treatment is perfectly valid with no obvious defects. Clearly, this one simulation does not represent all RHD flare simulations, but we have empirically found this to be true when testing multiple timesteps from the simulations used in the previous sections.

## 5.6 Case Study: Partial Frequency Redistribution

As highlighted by [Brown et al. \(2018\)](#), when considering the hydrogen Lyman lines it would be a significant improvement to simultaneously handle full time-dependence and PRD. It is possible that this could prove important for the energy balance of the line-forming regions, in a similar manner to the photoionisation effects discussed previously. Indeed, any significant variation in the hydrogen Lyman series may affect the populations of calcium and other species through photoionisation.  $H\alpha$  could also be affected through modifications of the level populations associated with the Lyman lines, as discussed by [Leenaarts et al. \(2012a\)](#). They investigated the same Doppler-like PRD approximation that was used in our F9 and F10 RADYN simulations presented in Sec. 5.4 and found that it was sufficient to obtain correct  $H\alpha$  line profiles in the quiet Sun case. The version of RADYN used by [Brown et al. \(2018\)](#) applied this same Doppler-like approximation. We term this treatment “Doppler-like” as it consists of the removal of natural radiative and van der Waals broadening terms, but the linear Stark broadening terms, that can be important in flares (e.g. [De Feiter & Švestka 1975](#)), are left in the model.

### 5.6.1 Modifying Our *Lightweaver*-based Model to Support PRD

The *Lightweaver* framework supports both time-dependence and PRD lines, thus we can attempt to modify the tool built in this chapter to also include PRD effects. [Hong et al. \(2019\)](#) included some of the effects of partial redistribution with simultaneous time-dependent populations by extracting the hydrogen level populations from RADYN and computing the PRD line profile ratio with RH. Using this method, the modified radiation field and line emission profile will not be taken into account in the calculation of the radiative rates used to update the populations. In the following, we attempt to treat this problem self-consistently.

Our hydrogen model atom is modified by restoring the natural and van der Waals broadening terms in the definitions of the  $Ly\alpha$  and  $Ly\beta$  spectral line models. These are the terms that were previously removed to render the line profiles Doppler-like. We then solve the statistical equilibrium problem for the initial atmosphere in PRD, and then update the PRD line emission ratio  $\rho_{ij}$  during each timestep based on the current atmospheric parameters and radiation field, performing several updates of  $\rho_{ij}$  and formal solutions with fixed populations, following the usual process described in Sec. 2.4.12. Unfortunately, this basic approach proved to have incredibly poor convergence. Many timesteps failed

## 5.6 Case Study: Partial Frequency Redistribution

to converge or the iterative scheme would enter a loop of repeatedly passing through the same states, getting no closer to the convergence threshold. Of the timesteps that do converge, many take an exceedingly large number of iterations.

These convergence problems are somewhat unexpected due to the reliability of the methods used here when applied to statistical equilibrium problems (e.g. [Uitenbroek 2001](#)). Unfortunately, we are faced with an additional problem when performing time-dependent simulations, especially in the reprocessing context used here: when a timestep fails to converge we lose the ability to investigate later in the simulation, as the populations at every timestep are dependent on those preceding. When reprocessing snapshots of an RHD simulation using a statistical equilibrium treatment, a failure for one of these to converge does not limit the analysis of any other snapshot.

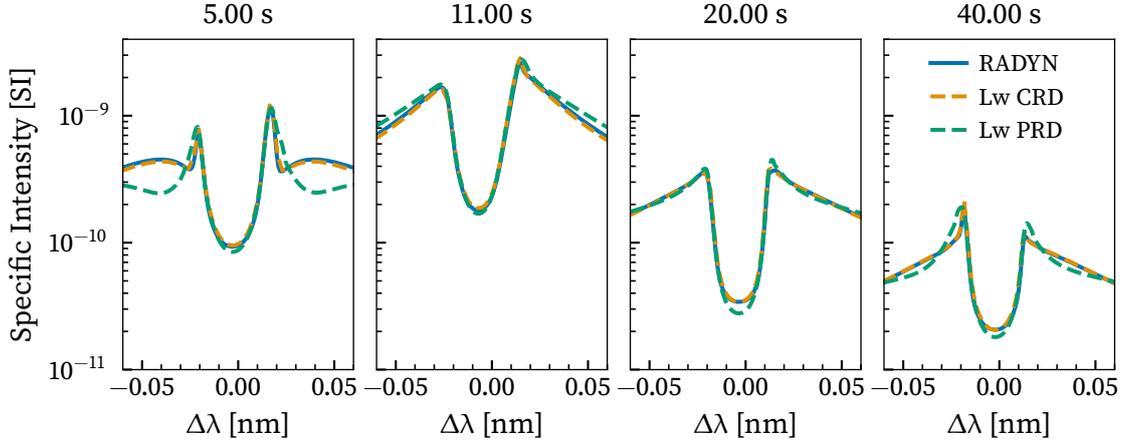
The following procedure was found to be relatively robust, and was sufficient to reprocess the complete F9 simulation presented previously. At the start of each timestep, when the populations are advected from the previous grid to the next, the line emission ratio  $\rho_{ij}$  is interpolated onto the new grid. Inherent to this assumption is the idea that  $\rho_{ij}$  will not change too drastically between subsequent timesteps. The time-dependent problem is then solved to our normal tolerance whilst keeping  $\rho_{ij}$  fixed at its interpolated values. Once converged, we continue to compute the time-dependent population update, whilst now interleaving iterations of updating  $\rho_{ij}$  following the usual process.

In the F9 simulation this approach proves to be rapidly convergent at early times, but far less so at later times during the cooling phase. Initially this appears unintuitive, but the non-thermal collisional rates from the electron beam depositing energy in the simulation are combined with thermal collisional rates during the heating phase and likely reduce the effects of PRD on these lines. At later times the Lyman line forming regions are likely to be less collisional, whilst the line remains strong with the radiative rates representing a significant component of the population transition rate. Similar convergence behaviour occurred when reprocessing the F10 simulation, but the convergence deteriorates sooner and despite adjustments to the number of PRD sub-iterations taken, the model usually fails to converge around  $t = 3.75$  s.

### 5.6.2 Results

The Ly $\alpha$  line computed in the reprocessed F9 simulation, both with approximate (dashed orange) and fully modelled PRD effects (dashed green), is shown in Fig. 5.12. It is compared

## 5.6 Case Study: Partial Frequency Redistribution



**Figure 5.12:** Comparison of the RADYN CRD, *Lightweaver* CRD, both with reduced broadening to approximate PRD, and *Lightweaver* PRD treatments of the Ly $\alpha$  line in the F9 model. Lw is used as shorthand for *Lightweaver*.

against the RADYN solution (blue), using the same approximate treatment. Once again, our CRD treatment of the problem with *Lightweaver* yields very good results when compared to RADYN as a reference solution. Other than the line wings at  $t = 5$  s and to a lesser extent at  $t = 40$  s the PRD and CRD (with reduced broadening terms) match well, although the PRD line core at  $t = 20$  s is a fair amount deeper than the CRD case. This is encouraging for the use of this approximation, especially in light of the difficulties converging the PRD problem in time-dependent simulations. Furthermore, the PRD solution, due to its poor convergence at later times, is extremely computationally costly when compared to the CRD solution, taking close to two orders of magnitude more CPU time (73 vs 1.2 wall hours<sup>6</sup> on *hephaistos*).

Ly $\alpha$  is not the only line affected by PRD in this simulation; we also investigated the effects on Ly $\beta$  and the Ca II H and K resonance lines. Ly $\beta$  was found to have very good agreement between the approximate and complete PRD treatments in *Lightweaver*, similar or better than that found for Ly $\alpha$ . The agreement between RADYN and the *Lightweaver* CRD treatment was also extremely good, as in the Ly $\alpha$  case. Ca II H and K are affected by PRD effects but also by the use of the LI treatment, albeit less than Ca II 854.2 nm, and this is likely due to the lack of overlap between Ly $\alpha$  and the Ca II resonance continuum. In Fig. 5.13, the variation of the Ca II K line computed using CRD with the LI treatment and

<sup>6</sup>Wall hours refer to a real world elapsed time, despite the fact multiple CPU cores are used in the problem. This metric also captures the fact that the implementation of updating  $\rho_{ij}$  is less parallelised in *Lightweaver* than the CRD MALI method which dominates the runtime of the CRD case (average  $\sim 65\%$  utilisation of 12 threads, rather than 98%).

## 5.6 Case Study: Partial Frequency Redistribution

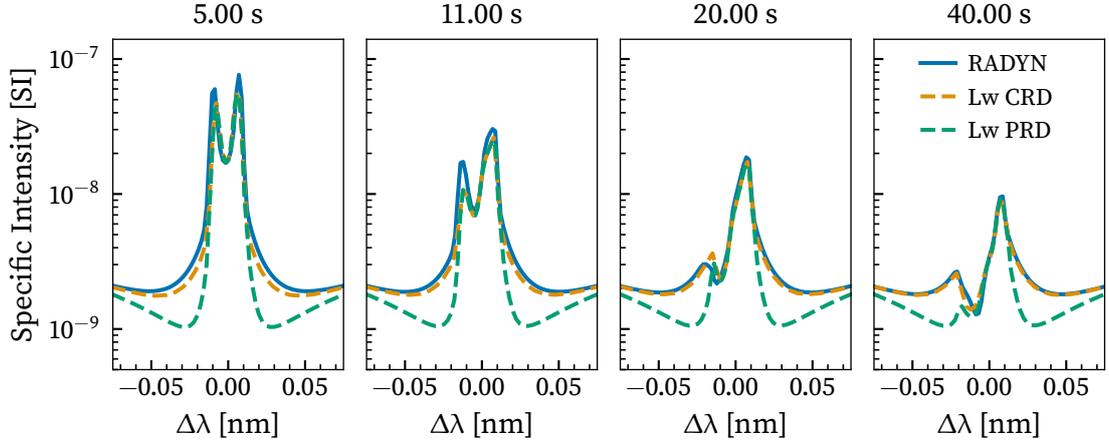


Figure 5.13: Comparison of the CRD and PRD treatments of the Ca II K line in the F9 model.

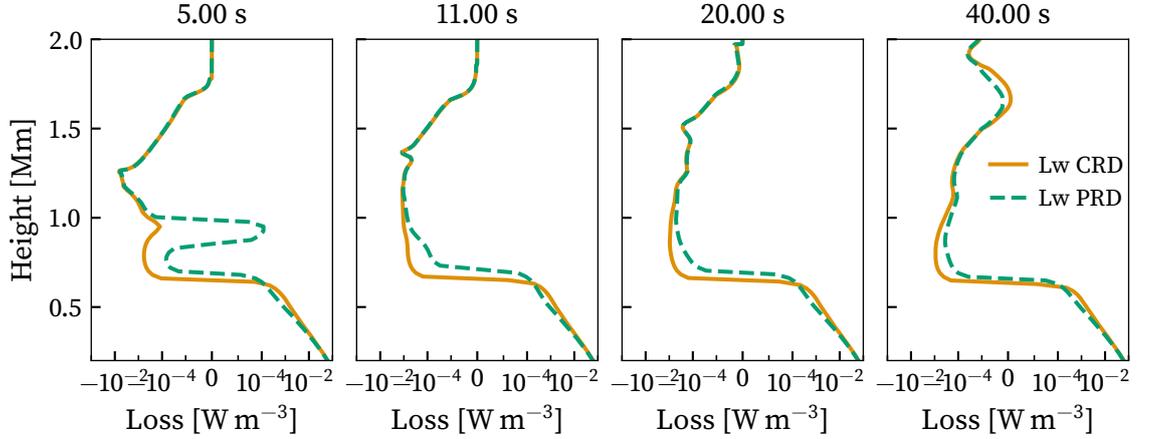
PRD with the LI treatment is shown. The differences between the CRD line profiles and the RADYN line profiles are due to the LI treatment. The PRD line cores agree very well with the CRD LI treatment, but as expected when a significant proportion of the scattering is coherent the line wings deepen and then slowly return to the same continuum level as the CRD case.

Uitenbroek (2002) undertook similar tests in a wave-heated quiet Sun RADYN simulation and looked at the radiative losses in the Ca II K line when treated in PRD (under statistical equilibrium). In the chromosphere these losses were found to vary between the CRD and PRD treatments by a factor of 2–5. Qualitatively, there is a smaller difference in the Ca II wings in our flare simulation than their wave-heated model. The losses between these two different treatments in the chromosphere are shown in Fig. 5.14. The differences in the radiative losses are most significant at  $t = 5$  s with the inversion of the sign of the losses between the PRD and CRD treatments around 1 Mm. Whilst less significant, at the other times shown there is also clear variation in radiative losses due to the treatment of the Ca II resonance lines in PRD, typically decreasing these relative to their CRD values.

### 5.6.3 Discussion

Whilst we have not attempted to implement this, it is possible that applying a direct, rather than iterative solution for evaluating the emission profile ratio may be more stable in this situation. The iterative method, implemented as first presented in Uitenbroek (2001), tends to have very good convergence when used in the statistical equilibrium

## 5.6 Case Study: Partial Frequency Redistribution



**Figure 5.14:** Comparison of the chromospheric radiative losses under CRD and PRD treatments of the Ca II K line in the F9 model. A negative value indicates a radiative loss and a positive value a radiative gain. Lw is used as a shorthand for *Lightweaver*.

case, whilst being less computationally expensive and much easier to apply in multi-dimensional geometry. It is plausible here that the effective damping of the population update equations of the time-dependent kinetic equilibrium equations relative to statistical equilibrium is reducing the convergence rate, due to the slower rate (per iteration) at which the populations evolve. It may therefore be worth investigating the direct solution in the hope that by more rapidly converging to the correct solution for the emission profile the populations do not pass through intermediate states that are not “valid” i.e. states that are intermediate to the iteration scheme, but do not occur at any point in the real world system we are computing. It is plausible that many of the intermediate states that occur in both time-dependent and statistical equilibrium calculations are not states that occur naturally, instead resulting from our numerical treatment of the system.

When additional iteration processes such as PRD and charge conservation are applied (leading to the application of three simultaneous iterative processes to the problem), whilst being combined with the damped nature of the time-dependent population update, it appears that cycles in the global iteration procedure can occur quite frequently, and no method that we have found reliably prevents these (including interleaving basic  $\Lambda$ -iterations and applying Ng acceleration (Ng 1974)). Avrett & Loeser (2008) also comment that strong lines can pose problems for the convergence of ALI schemes and propose a scheme for computing strong PRD lines by solving a linear system of simultaneous equations for the rate and transfer equations, whilst solving for weaker transitions by an ALI method. If this approach is reformulated and implemented for the time-dependent case, it is possible that it will not suffer the same convergence problems that were encountered

in this section, especially given the strength of these lines in flare-heated atmospheres.

## 5.7 Digging Deeper: Time-Dependent Response Functions

In Section 5.4.2 the contribution function was used to investigate the effects of the LI treatment on the Ca II 854.2 nm line. It may be possible to obtain greater interpretability through the use of response functions (as introduced in Sec. 3.2.2), but to our knowledge these have only been considered for models in statistical equilibrium. Carlsson & Stein (2002) investigated the settling times for hydrogen ionisation in RADYN simulations by applying a perturbative numerical formulation similar to that we will apply in computing time-dependent response function.

For the time-dependent situation, we perturb an atmospheric parameter at a particular depth as before, and then a timestep by which to advance the system is needed. This should be short compared to the hydrodynamic evolution of our system, but sufficiently long to allow the population change to rise above our convergence threshold and correctly converge to the new solution. In essence, we are computing the time-dependent response function  $\mathcal{R}_{\nu,q,t}$  as the response function  $\mathcal{R}_{\nu,q}$  after a time  $\Delta t$  has elapsed. As discussed in Sec. 3.2.2, a centred finite-difference method in the perturbed parameter is applied, i.e.

$$\begin{aligned} \mathcal{R}_{\nu,q,t}(k, t) &= \mathcal{R}_{\nu,q}(k) \Big|_{t+\Delta t} \\ &= \frac{I(\nu, q_k + \delta q/2, t + \Delta t) - I(\nu, q_k - \delta q/2, t + \Delta t)}{\delta q}. \end{aligned} \quad (5.3)$$

Carlsson & Stein (2002) and Judge (2005) note that for a two level atom a population perturbed out of statistical equilibrium will follow an exponential decay back to the steady state solution, further supporting the choice of a short timestep to attempt to remain within the more “linear” regime of this process. The NLTE multilevel system is more complex and theoretically depends on the eigenvalue spectrum of the rate matrix. In the analysis of Carlsson & Stein (2002) the full rate matrix was found to significantly underestimate the ionisation relaxation time, and NLTE effects slowed the population evolution. Judge (2005) applied a simplified variant of the mathematical framework of Gayley (1990) showing how, at high optical depths, NLTE radiative transfer effects can reduce the radiative decay rate. Extreme care should be taken when trying to make quantitative use of these time-dependent response functions, especially for longer timesteps, but for now we shall focus on qualitative aspects.

### 5.7 Digging Deeper: Time-Dependent Response Functions

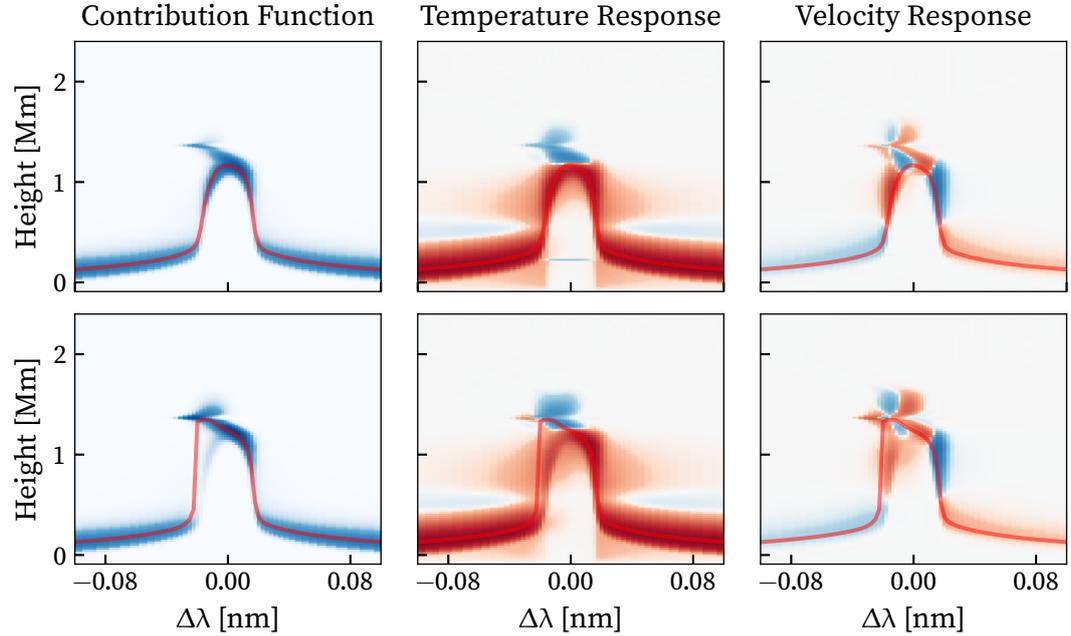
Time-dependent response functions were computed for perturbations in temperature and velocity to the same atmospheres for which the contribution functions were considered in Section 5.4.2. These properties were chosen as they are important but “relatively” free parameters in the system. It is also common to compute the response functions to electron density, microturbulent velocity and magnetic field, however these parameters are either not considered or not truly free. Following RADYN, the microturbulence is assumed to be constant both in time and throughout the atmosphere, and this model does not consider any effects of the magnetic field. For the LI model, the electron density is dependent on the rest of the atmosphere through charge conservation, and is therefore not a free parameter. In the LE case, where the hydrogen populations are considered fixed over each timestep, the electron density is also held fixed over the duration of the response functions. We note that the radiative response to the electron density could likely be computed following [Metcalf et al. \(1990a\)](#) who present the following method for computing the response to electron density  $n_e$

$$\frac{\partial I}{\partial n_e} = \left( \frac{dI}{dT} - \frac{\partial I}{\partial T} \right) \left( \frac{dn_e}{dT} \right)^{-1}, \quad (5.4)$$

where these terms are defined as follows: the total derivative of  $I$  with respect to  $T$  can be evaluated from computing the response with charge conservation enabled, the partial derivative of  $I$  with respect to  $T$  from computing the response with  $n_e$  fixed, and the derivative of  $n_e$  with respect to  $T$  can simply be computed from the change in  $n_e$  during evaluation of the response function with charge conservation. This method was originally presented for computing the response of an intermediate parameter such as the source function or opacity to  $n_e$ , but there is no reason why it cannot be applied directly to intensity. In the context of inversions, this additional response is very useful, however our RHD models should normally remain self-consistent. Nevertheless, this could also be useful for interpretation, but is  $\sim 2\times$  more computationally costly than other parameters due to the requirement of additional derivatives.

Our contribution and response functions are shown in Fig. 5.15 – 5.18 (LI in the upper rows and LE in the lower). In the response function plots red shows an enhancement in outgoing radiation as a response to an increase in the perturbed parameter at the particular depth, whereas blue indicates a reduction. The same colour scale is used for all response function panels. All of the response functions shown in these plots were computed with a timestep of 1 ms, but were visually identical to those computed for a timestep of 10 ms and extremely similar to those computed for a timestep of 0.1 ms. Substantial differences appeared for a timestep of 1  $\mu$ s, possibly due to convergence issues, as the populations have little time to evolve over such a short timestep, making it difficult

### 5.7 Digging Deeper: Time-Dependent Response Functions



**Figure 5.15:** Comparative plots of the temperature and velocity response functions and the contribution function for the Ca II 854.2 nm line in the F9 simulation at  $t = 11$  s. The upper row shows the LI treatment, and the lower row the LE treatment. For the response functions, enhancement to a positive perturbation is shown in red, and reduction in blue. The same colour scale is used in all response function panels, and  $\tau_v = 1$  altitude is overlaid on each plot.

to judge convergence to the necessary level. The stability of the response across these three timesteps suggests we are capturing a physical response, rather than the numerics of the method.

The shape of the temperature response function is immediately recognisable as similar to the contribution function shown in the first column. This is not surprising, as temperature is the most important atmospheric parameter in spectral line formation, and will always affect the region where the observed photons form. Nevertheless, at a particular wavelength, temperature response features can be seen at greater depth (well below the  $\tau_v = 1$  line) than any in the contribution function, as a change in radiation field at this depth can affect the radiative rates in the line-forming region.

The other immediately remarkable feature present in all of the temperature response functions shown here is the presence of a negative response to temperature above  $\sim 1.2$  Mm (i.e. an increase in temperature in this region decreases the outgoing intensity). This is due to a reduction in the local source function, primarily due to an increase in the

5.7 Digging Deeper: Time-Dependent Response Functions

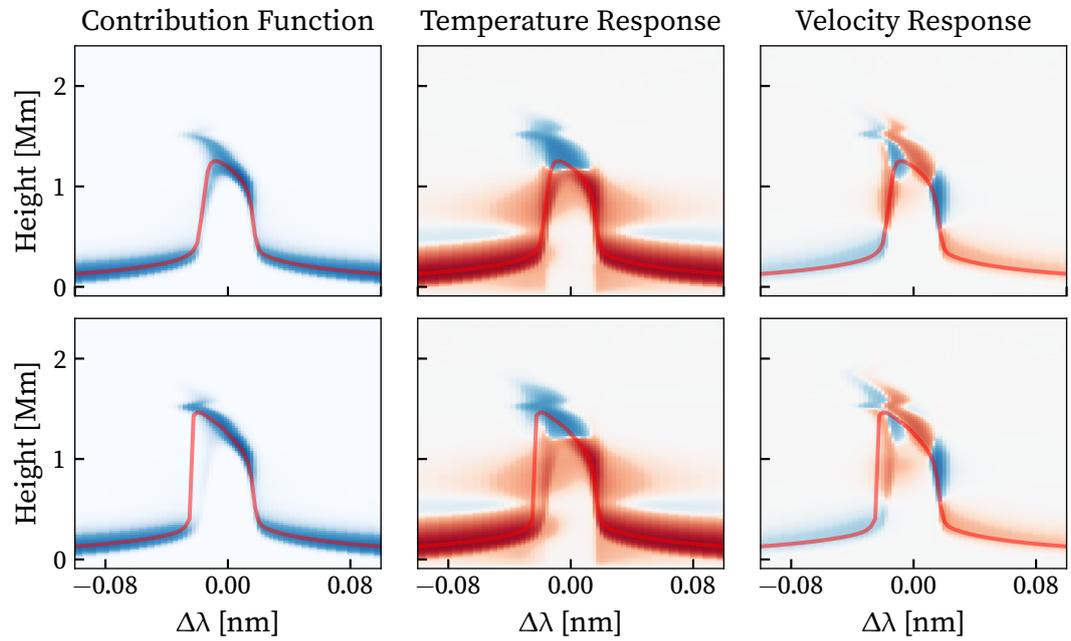


Figure 5.16: Comparative plots of the response and contribution functions, equivalent to Fig. 5.15 for the F9 simulation at  $t = 20$  s.

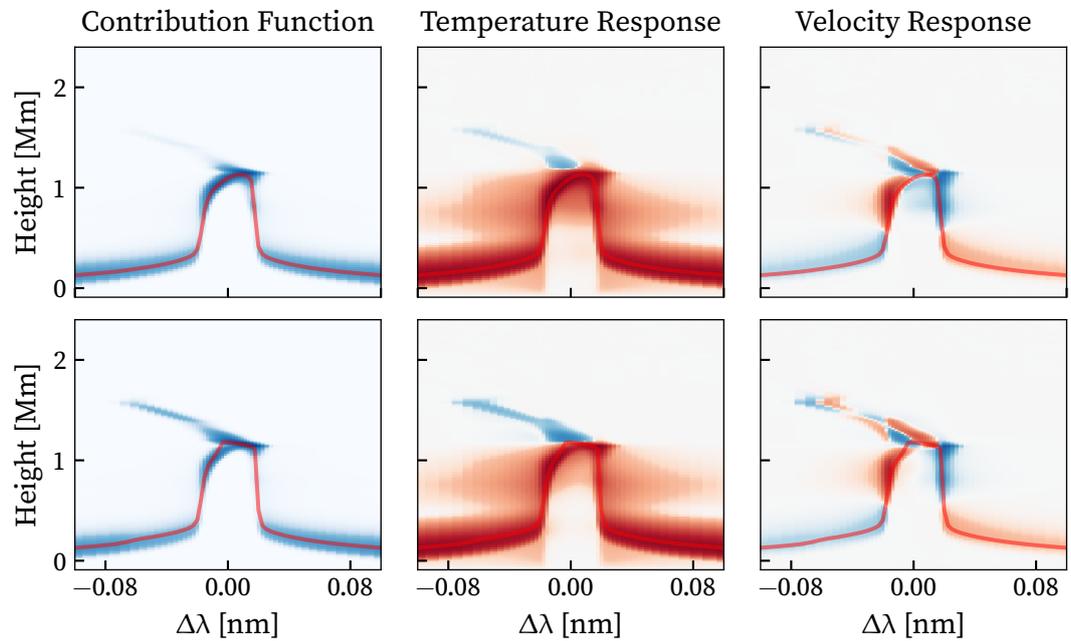
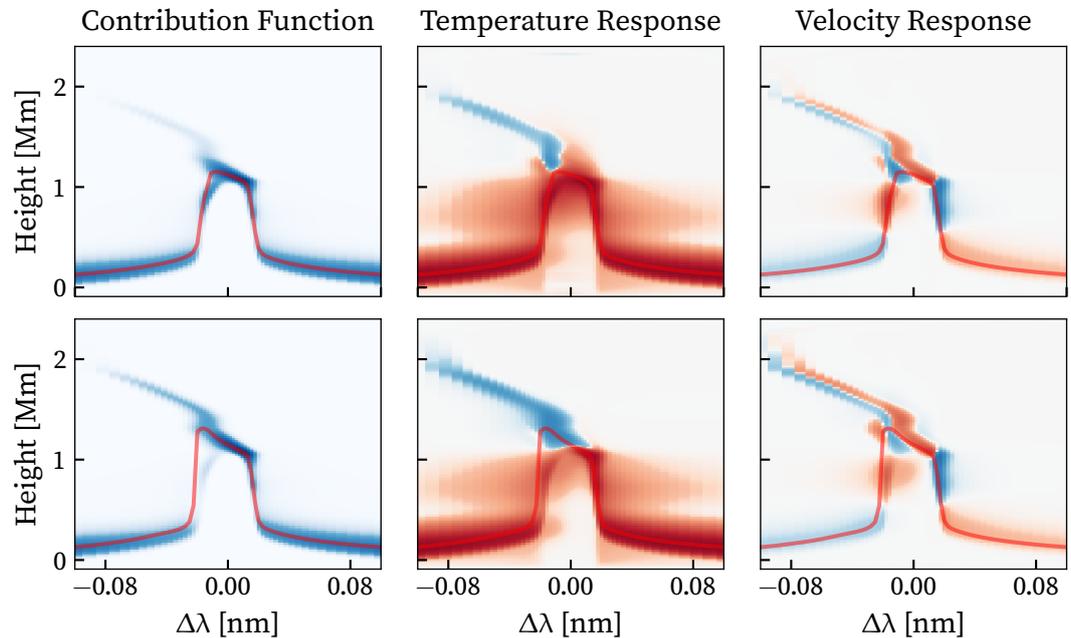


Figure 5.17: Comparative plots of the response and contribution functions, equivalent to Fig. 5.15 for the F10 simulation at  $t = 11$  s.

### 5.7 Digging Deeper: Time-Dependent Response Functions



**Figure 5.18:** Comparative plots of the response and contribution functions, equivalent to Fig. 5.15 for the F10 simulation at  $t = 20$  s.

populations of the lower level of this transition: below 1.4 Mm this reduces the strength of the line core, whereas above this region this reduction is due to an increase in opacity, trapping radiation in the atmosphere, and creating a dip in the line profile. This same region is simply shown as contributing in the contribution function panels and tends to be more pronounced in models with the LE treatment, where the  $\tau_v = 1$  line is at a higher altitude.

There is a narrow blue bar in the line core of the LI temperature response function of Fig. 5.15 around 0.25 Mm. This is likely an artefact of the numerical method used to compute the response function, although it also appears for the other timesteps used to verify the response function. When a feature like this appears only for one depth point it is usually a spurious numerical artefact, and we will discount it in the following.

In the right-most column of these figures we have plotted the velocity response function. The structure of this plot below  $\sim 1$  Mm is easy to interpret; an increase in velocity shifts the line-core towards bluer wavelengths, and away from redder wavelengths. The effects on the far wings, visible below  $\sim 0.4$  Mm are less immediately intuitive, but are very small variations corresponding to increased absorption of photospheric emission in the blue wing, and reduced absorption in the red wing, due to the far wings of the line profile. The

### 5.7 Digging Deeper: Time-Dependent Response Functions

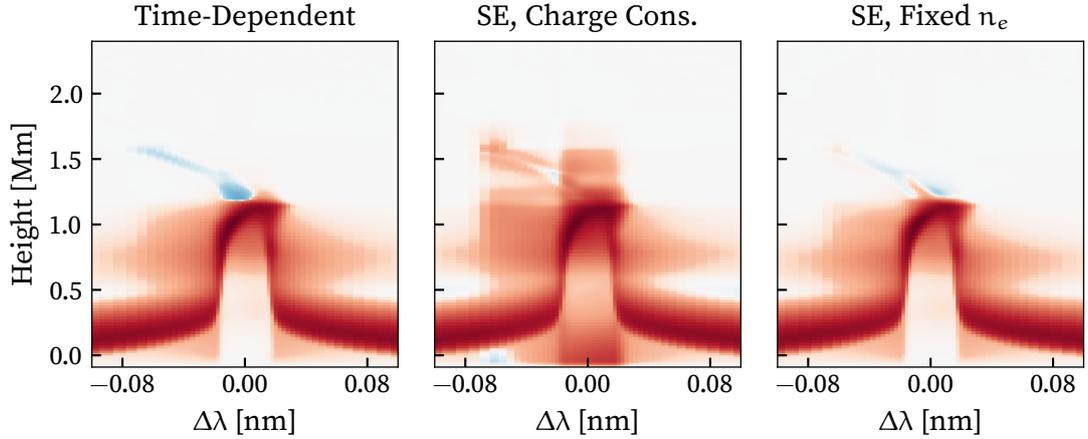
structure in the line core is far more complex, and has a non-trivial response to variations in velocity, affecting the LE and LI cases differently. This is most pronounced in the F10 response function at 11 s (Fig. 5.17). There are also responses due to changes in velocity from well below the  $\tau_v = 1$  level in the line core which produce changes in opacity in these regions affecting the radiation field in the line forming region.

The difficulties of a quantitative analysis become clear when looking at the scales of two different LI temperature response functions from the F10 simulation at  $t = 11$  s. Normalising the response to the line profile, for a timestep of 10 ms the maximum response in the Ca II 854.2 nm line was  $8.84 \times 10^{-5} \text{ K}^{-1}$ , and for a timestep of 0.1 ms this was  $8.86 \times 10^{-5} \text{ K}^{-1}$ . The consistency of these values and the response functions suggests that the evolution is occurring on a timescale  $\leq 0.1$  ms. This is not overly surprising given the conclusions of Sec. 5.5, where a statistical equilibrium treatment was found to adequately reproduce the full time-dependent evolution of the calcium populations at most points in the model presented. This consistency poses problems to recasting these response functions as a fractional change in the line profile per unit time, as the model with the shorter timestep would appear to have a response two orders of magnitude larger than the smaller one.

More interestingly, the  $\text{H}\alpha$  response function computed for these two timesteps was also found to be almost identical, despite the accepted need for a time-dependent treatment. It is probable that we are seeing results similar to those of Leenaarts et al. (2012a), who found that for RMHD simulations  $\text{H}\alpha$  can be synthesised correctly in statistical equilibrium if the equation of state takes into account non-equilibrium ionisation.

This leads to the question of whether it is possible to employ statistical equilibrium response functions in these simulations. These are compared both with and without charge conservation to the time-dependent temperature response function for the Ca II 854.2 nm line in the F10 atmosphere at  $t = 11$  s in Fig. 5.19. The left-hand panel once again shows the time-dependent response function computed with a timestep of 1 ms, the centre panel shows a statistical equilibrium response function, including hydrogen evolution and charge conservation, while the right-hand panel shows a statistical equilibrium response function computed whilst holding the hydrogen populations and electron density fixed. Immediately we see a large difference between the middle panel and the other two. This is due to the hydrogen populations, and more specifically its ionisation state, relaxing back to the statistical equilibrium solution. The information presented on this middle panel is not equivalent to the left-hand panel, showing responses in the line core at locations both deeper and shallower than those found in the time-dependent solution. The right-hand

### 5.7 Digging Deeper: Time-Dependent Response Functions



**Figure 5.19:** Comparison of time-dependent and statistical equilibrium response functions for the Ca II 854.2 nm line in the F10 simulation at  $t = 11$  s. The time-dependent response function is computed with a timestep of 1 ms. The middle column shows the statistical equilibrium response function when charge conservation is considered, and the right-hand column the statistical equilibrium response function when the hydrogen populations and electron density are held fixed.

panel (holding the hydrogen populations and electron density fixed) agrees better with the time-dependent response function, but there are still significant differences in the structure above 1.2 Mm, with additional positive responses appearing. It was found that running the time-dependent response functions for long timesteps ( $> 100$  s), started to produce similar structure, although these did not converge to the same result. It may be acceptable to use this second form of statistical equilibrium response function to diagnose the Ca II 854.2 nm response in RHD simulations, but it typically took longer to converge than the time-dependent approach, and cannot be generalised to compute response functions of hydrogen lines, due to these populations being held fixed. The difference between the time-dependent and statistical equilibrium with fixed electron density response functions is qualitatively similar in the F9 simulation and at other times in this simulation, although at later times ( $t \gtrsim 35$  s), as the populations settle the differences are reduced in both simulations.

The time-dependent response functions discussed here are a more NLTE-motivated approach to the problem of spectral line formation in RHD simulations. They allow the identification of the regions and parameters to which the line is sensitive (which may be due to non-local radiative effects), and show whether a change in a parameter will provoke a positive or negative response on the outgoing radiation. This differentiates response functions from the contribution function by showing the negative response to

temperature above  $\sim 1.2$  Mm. Response functions can also be used to separate the effects of different parameters in these models, a feature that will likely become increasingly important as the complexity of flare modelling increases further. We stress that as response functions represent a local gradient, they are only valid in small parameter regions around those the model was computed with. The time-dependent formalism presented here will need additional development for quantitative uses, but we feel that it already provides information complementary to the contribution function and are optimistic about future applications to improve the interpretation of line formation in RHD simulations

## 5.8 Discussions

We have presented multiple applications of the *Lightweaver* framework to reprocessing the radiative transfer of RADYN simulations. Through the construction of simple tools, we are able to quantitatively investigate the importance of various radiative effects that may not be considered in RADYN's treatment (but without modifying the energy balance or hydrodynamics), more rapidly than would be possible incorporating the necessary changes into RADYN. These tools have been used to investigate the effects of photoionisation of Ca II to Ca III by the hydrogen Lyman lines, the importance of treating the Ca II populations with time-dependence, and the possible difficulties in applying a PRD treatment in RHD simulations.

From the flare simulations presented here the hydrogen Lyman lines have significant effects on the Ca II photoionising radiation field. This results in substantial changes in both the outgoing Ca II 854.2 nm line profiles, and the radiative energy balance in the upper chromosphere. The radiative losses in this region were found to be affected by up to 15%, which could in turn lead to differences in the hydrodynamic evolution of the simulation and thus greater difference in the line profiles. We did not find a clear correlation between the magnitude of this effect and the increased heating in the F9 and F10 models: in both cases the effect was important, and therefore is likely to remain so at higher energy depositions. It is therefore essential that RHD simulations start to consider this in a self-consistent way to gauge its full effect. These photoionisation effects could also affect other species with continua overlapping the hydrogen Lyman lines. For instance, Mg II line profiles are often used for chromospheric diagnostics, but these are unlikely to be as significantly affected by the hydrogen Lyman lines as Ca II due to the resonance continuum edge being located at 82.46 nm, which can therefore only be affected by the far Lyman continuum. The Mg II populations will be influenced to a lesser extent

## 5.8 Discussions

by photoionisation from the subordinate continua, some of which overlap the Lyman series.

On the other hand, we found that the deviations of the calcium populations due to full time-dependent kinetic equilibrium over statistical equilibrium were typically small in the simulations investigated. This is interesting, as it theoretically allows for Ca II to be processed separately, but as the radiative losses from this species need to be known for the hydrodynamics of an RHD simulation, this species likely still needs to be treated in lockstep with the rest of the simulation. Additionally, there is typically a loss in speed by computing the statistical equilibrium solution at each timestep in the simulation (even starting close to the solution from the previous timestep's solution) over the time-dependent solution when using traditional radiative transfer methods. This could change with the application of machine learning, as it will be far easier to construct detailed models that provide the unique statistical equilibrium solution for a given atmosphere than determining the correct time-dependent solution. As the hydrogen populations would continue to require a full time-dependent treatment there is not likely to be an immediate gain from such a model in the context of RHD simulations. However, the nuanced analysis of the formation of H $\alpha$  in the quiet Sun by [Leenaarts et al. \(2012a\)](#) suggests that it may be sufficient to purely capture the effects of hydrogen non-equilibrium ionisation (i.e. the fraction of hydrogen that is fully ionised and its effects on the electron density) through the equation of state and then perform NLTE synthesis in statistical equilibrium with this given electron density. This would also need to be validated for flare models, which could be done following the approach laid out in this chapter. It would then become significantly easier to apply machine learning models to these separable components; this could provide massive computational gains in RHD modelling, whilst detailed spectral synthesis would remain possible through traditional NLTE radiative transfer approaches. Parallels can be drawn between the application of machine learning to this problem and the use of simplified local rates following the method of [Sollum \(1999\)](#) as used in HYDRAD ([Reep et al. 2019](#)), and BIFROST ([Gudiksen et al. 2011](#); [Leenaarts et al. 2007](#)), but it remains to be proven whether this method retains sufficient accuracy at the high energy depositions that occur in flares.

Despite the apparent simplicity of incorporating PRD into these tools, the iteration process was plagued with convergence problems that we have been unable to fully resolve. Nevertheless, for the hydrogen Lyman lines in the F9 simulation presented in this chapter, the Doppler-like PRD approximations currently present in RADYN perform well, and come at no additional computational cost over the basic CRD treatment, validating the quiet Sun results of [Leenaarts et al. \(2012a\)](#). We therefore recommend that this approximation

## 5.8 Discussions

continue to be employed for Ly $\alpha$  and Ly $\beta$  until more advanced iteration methods are developed and implemented (that can simultaneously efficiently solve for the atomic level populations, electron density, and necessary PRD line emission profile ratios), although it would be good to test the quality of this approximation in more strongly heated atmospheres.

It is possible to apply the method outlined in Sec. 5.6 to Mg II, but it is likely, as shown by Kerr et al. (2019a), that treating it in statistical equilibrium with PRD will provide a sufficiently accurate solution whilst remaining relatively efficient, hopefully allowing this species to be incorporated into RHD simulations in the near future. It is also important to investigate other effects that may change the formation of these spectral lines. For example, Zhu et al. (2019), used an improved Stark broadening treatment when synthesising Mg II h & k from snapshots of a RADYN model in RH. They found that a 30 $\times$  increase in Stark broadening (or the equivalent increase in another Lorentzian broadening term) was needed to reproduce observations. Additionally, Kowalski et al. (2017a) note that a correct treatment of Mg II requires the modelling of overlapping bound-bound transitions, which is not possible with current RHD codes, but does not pose any great computational difficulty in theory.

The effects of treating the Ca II resonance lines in PRD were significantly larger than those seen with the hydrogen Lyman lines. This is seen not only in the line profiles, but also in the chromospheric radiative losses in these lines, which vary substantially when treated with PRD (and are typically smaller than those computed in CRD, as found previously in wave-heated atmospheres by Uitenbroek (2002)). It is therefore important to investigate whether applying PRD to the H and K lines can be done efficiently. Empirically, we have found that the majority of convergence problems originate from the hydrogen Lyman lines, thus it may be relatively efficient to treat the Ca II resonance lines in PRD. If similar convergence issues also affect these lines, then it would be desirable to devise and implement an approximation scheme similar to that used for Ly $\alpha$  and Ly $\beta$ . As previously mentioned, it has been suggested that RADYN's treatment of the Ca II resonance lines in CRD overestimates the radiative losses but this overestimation is compensated by not considering the Mg II h and k lines (e.g. Carlsson & Stein 2002; Kerr et al. 2019b). The techniques presented in this chapter should allow for a thorough investigation of this assumption and the development of suitable PRD approximations if necessary.

In summary, we found that it is essential to consider the photoionising effects of the Lyman lines when modelling the Ca II lines in flares, but a statistical equilibrium treatment of the Ca II level populations appears to be sufficient in general. Considering PRD

## 5.8 Discussions

effects, we found that the Doppler-like approximation used in RADYN appears accurate for  $\text{Ly}\alpha$  and  $\text{Ly}\beta$ , and similar approximations should be investigated for the Ca II resonance lines, as their line profiles and radiative losses vary dramatically between PRD and CRD treatments (as previously remarked by [Uitenbroek \(2002\)](#) in wave-heated chromospheric simulations).

# 6

## Two-Dimensional Radiative Transfer

It is clear that the world around us is multi-dimensional, but until this point we have only considered radiative transfer in one-dimensional plane-parallel atmospheres.

All of the theory of radiative transfer discussed in previous chapters remains valid when applied to higher dimensional systems, with the exception of the description of the formal solver. This is due to the non-local terms that appear within the MALI description (with diagonal  $\Lambda$  operator) being handled by the formal solver, which is responsible for computing the radiation field throughout the plasma from the local parameters and boundary conditions and thus coupling the atmospheric nodes to each other. Once the radiation field has been computed it is then used as a local parameter in the rest of the iteration. In fact, if the storage for the atmosphere is “flattened” into a one-dimensional form, the code from the plane-parallel case can (and should) be used to implement the iteration scheme.

In this chapter we shall first describe the extension of the *Lightweaver* framework to two dimensions (with the possibility of further extension), and then describe its application to the simulation a slab of quiet sun being illuminated by an adjacent slab of quiet sun, along with potential implications for future observations at high resolution.

### 6.1 The Formal Solver in Two-Dimensions

Similarly to the plane-parallel formal solver used in *Lightweaver*, described in Sec. 2.4.2, we use the short-characteristics method to compute the radiation field throughout the atmosphere. This approach was first employed in two-dimensions as a pure parabolic treatment by Kunasz & Auer (1988) and was developed further by Auer & Paletou (1994) to produce

## 6.1 The Formal Solver in Two-Dimensions

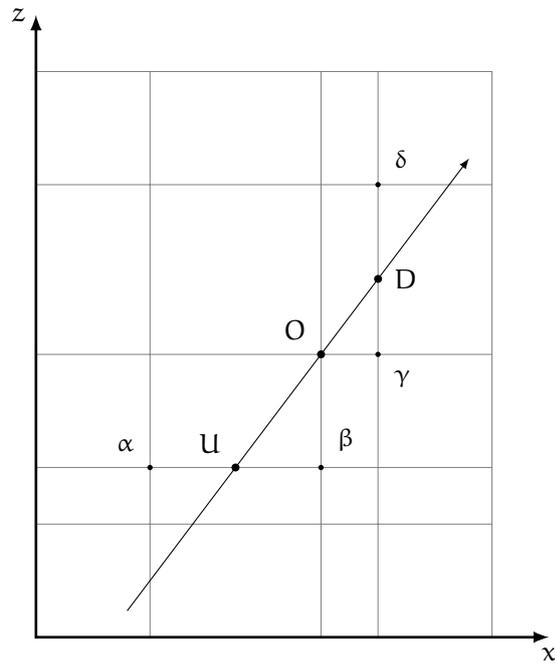


Figure 6.1: Diagram of short-characteristics formal solver in two-dimensions.

the aforementioned limited parabolic scheme that avoids possible over- and undershoots. We assume the following Cartesian basis: the  $z$ -axis is oriented as in the plane-parallel case, oriented vertically from photosphere to corona, the  $x$ -axis is perpendicular and co-planar to the  $z$ -axis (in the plane of the page for the following diagrams), and by the right-hand rule the  $y$ -axis is oriented into the plane of the page. In the two-dimensional case it is assumed that the atmospheric parameters are homogenous along the  $y$ -axis, but vary along the  $x$ - and  $z$ -axes. We assume that the atmosphere has a fixed stratification in  $x$  and  $z$ , and that the atmospheric parameters are known at each intersection of these grids.

The mean intensity at each point will be computed similarly to the plane-parallel case: by integration of the local intensity over a weighted angular quadrature. The Gauss-Legendre nodes used in the plane-parallel case are poorly suited to anisotropy that occurs in the two- and three-dimensional cases, and so we therefore employ the optimised angular quadratures of Štěpán et al. (2020) in *Lightweaver*.

For each ray prescribed by the angular quadrature the formal solver must perform one sweep through the grid. The general case, that of an inclined ray travelling through the atmosphere is shown in Fig. 6.1. In the case of this ray the formal solver must sweep first

## 6.1 The Formal Solver in Two-Dimensions

along  $x$  and then along  $z$ . The difficulties which can arise from this will be discussed in Sec. 6.1.2. The points  $U$  and  $D$  refer to them being “upwind” and “downwind” of the point  $O$  for which we are currently computing the intensity. This can be visualised by looking at the intersections of ray with the grid. To compute the intensity in the direction of this ray at point  $O$  using the short-characteristics formulation we have, as in one-dimension,

$$I_O = I_U e^{-(\tau_U - \tau_O)} + \int_{\tau_O}^{\tau_U} S(t) e^{-(t - \tau_O)} dt, \quad (6.1)$$

where these terms have their usual meanings,  $I_O$  is the intensity in this direction at  $O$ ,  $I_U$  is the intensity in this direction at  $U$ , and we have dropped the angular and frequency dependencies for clarity. Thus, to compute the intensity at  $O$ , the intensity at  $U$  must first be known. As  $U$  does not generally lie on a point of our discrete two-dimensional grid, the value of  $I_U$  (and other quantities) are not computed directly by the formal solver, and must instead be interpolated from the values of grid points along the line  $\alpha\beta$ .

Following the short-characteristic method, a functional form must be assigned to the variation of  $S$  over the the line segment  $[UO]$ . In the simplest case, this can again be a linear functional form, but due to the additional dimension for inhomogeneities in the two-dimensional case, unless the grid is very fine or the atmosphere very slowly varying, this may be a poor choice. A higher order parametrisation of  $S$  is likely to require the values of the source function at both  $U$  and  $D$ , and possibly other points along the ray. Similarly to the cubic Bézier spline<sup>1</sup> method used as standard in our plane-parallel code, we once again turn to monotonic Bézier splines for safe, smooth interpolation, minimising the presence of under- and over-shoots. The cubic method we apply in plane-parallel atmospheres requires four points along  $UD$ , which becomes less practical in higher dimensions, due to the computational demands of the method. Instead we choose the quadratic Bézier spline method, BESSER, of Štěpán & Trujillo Bueno (2013), which will be briefly summarised here for the scalar case of the RTE.

### 6.1.1 The BESSER method

The BESSER method (Štěpán & Trujillo Bueno 2013) differs from other monotonic Bézier spline methods by ensuring the continuity of the first derivative of the interpolant at  $O$ . Due to the large differences in optical depth between adjacent regions (e.g.  $\tau_{UO}$  and  $\tau_{OD}$ )

---

<sup>1</sup>A superb graphical introduction to Bézier spline methods and a visualisation of their construction and features is provided in the video by Freya Holmér, *The Beauty of Bézier Curves* available at <https://youtu.be/aVwxzDHniEw>.

### 6.1 The Formal Solver in Two-Dimensions

that are likely to occur in multi-dimensional cases with irregular grids, this method is designed so as to guarantee that if the values over [UOD] are monotonic then the spline interpolant will remain monotonic.

The spline interpolant over the [UO] interval is described by

$$f(a) = (1 - a)^2 f_U + 2a(1 - a) c_U + a^2 f_O, \quad a \in [0, 1], \quad (6.2)$$

where  $f_U$  and  $f_O$  are the values of  $f$  at points U and O respectively,  $c_U$  is the functional value of the spline's control point, and  $a$  is the normalised coordinate for the distance along [UO]. The control points are points half way along an interval, defining the tangent to the spline at each end of the range i.e.  $c_U O$  defines the tangent to the spline at O and  $U c_U$  defines the tangent to the spline at U. If we denote the coordinates of our points along the ray  $s_U, s_O$ , and  $s_D$  then we have  $a = (s - s_U)/h_U$ , where  $h_U = s_U - s_O$  and  $s \in [s_U, s_O]$ . An equivalent interpolation can be defined over the [OD] segment. The monotonicity (for monotonic  $f_U, f_O, f_D$ ) and continuous first derivative of the interpolating functions at O is ensured by following this procedure:

1. Verify the monotonicity of  $f_U, f_O, f_D$ , and if  $f_O$  is a local extremum then the control points  $c_U$  and  $c_D$  are set to  $f_O$ , giving a derivative of zero at O. As this is the only possible solution for this case, the process stops here.
2. Compute an estimate of the derivative at O, by using the derivative of the standard parabolic interpolation of UOD.
3. Use this derivative to compute the initial values at the control points, by direct projection of the first derivative to their coordinates (as this defines a tangent to function at O).
4. Check  $c_U \in [f_U, f_O]$ . If not, set  $c_U$  to  $f_U$  to correct for any new extremum and halt the process, as  $c_D$  is not needed for the integration of the source function.
5. Check  $c_D \in [f_O, f_D]$ . If not, set  $c_D$  to  $f_D$ , and use this to compute a new value for the derivative at O, and follow the projection of this tangent to determine  $c_U$  (due to the enforced continuity of the derivative at O).

This process is described in more detail in Štěpán & Trujillo Bueno (2013), but this covers the most important elements of the process.

### 6.1 The Formal Solver in Two-Dimensions

Similarly to the plane-parallel case, we will solve the RTE in optical depth, due to its increased stability. This interpolation method can be used to compute the optical depths  $\tau_{\text{UO}}$  and  $\tau_{\text{OD}}$  along these segments by

$$\tau_{\text{UO}} = \frac{1}{3}(\chi_{\text{U}} + \chi_{\text{O}} + \chi_{c_{\text{U}}})(s_{\text{O}} - s_{\text{U}}), \quad (6.3)$$

$$\tau_{\text{OD}} = \frac{1}{2}(\chi_{\text{O}} + \chi_{\text{D}})(s_{\text{D}} - s_{\text{O}}). \quad (6.4)$$

Note that a linear approximation was used for  $\tau_{\text{OD}}$ , as the previous process does not guarantee the calculation of  $c_{\text{D}}$ . This could be modified to also use the quadratic spline method, which may be more robust, however the value of  $\tau_{\text{OD}}$  rarely affects the final solution dramatically.

This method can now be applied to the source function but parametrised along  $\tau$  rather than geometric distance along the ray  $s$ . The BESSER quadratic spline method is then used to compute the value of the source function control point  $S_{c_{\text{U}}}$ . The integral in (6.1) can now be evaluated, by using the prescribed quadratic spline variation of  $S$  over the interval  $[\tau_{\text{U}}, \tau_{\text{O}}]$ . Working through the maths we arrive at

$$I_{\text{D}} = I_{\text{U}}e^{-\tau_{\text{UO}}} + \omega_{\text{U}}S_{\text{U}} + \omega_{\text{O}}S_{\text{O}} + \omega_{c_{\text{U}}}S_{c_{\text{U}}}, \quad (6.5)$$

where,

$$\omega_{\text{U}} = \frac{2 - (\tau_{\text{UO}}^2 + 2\tau_{\text{UO}} + 2)e^{-\tau_{\text{UO}}}}{\tau_{\text{UO}}^2}, \quad (6.6)$$

$$\omega_{\text{O}} = 2 \frac{\tau_{\text{UO}} - 2 + e^{-\tau_{\text{UO}}}(\tau_{\text{UO}} + 2)}{\tau_{\text{UO}}^2}, \quad (6.7)$$

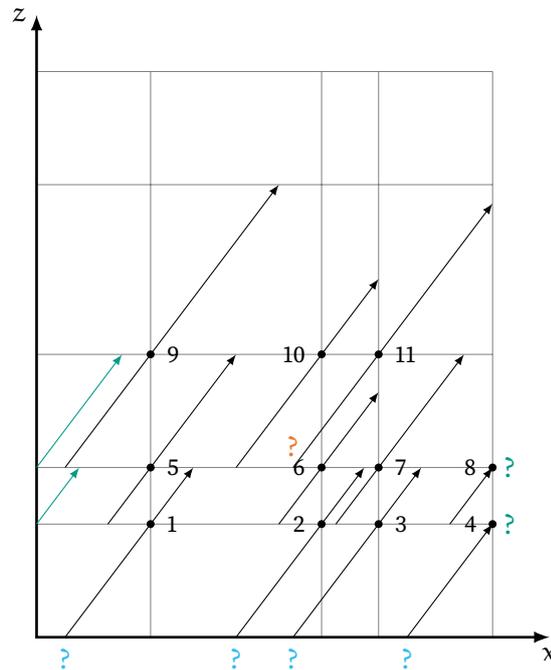
$$\omega_{c_{\text{U}}} = 1 - 2 \frac{e^{-\tau_{\text{UO}}} + \tau_{\text{UO}} - 1}{\tau_{\text{UO}}^2}. \quad (6.8)$$

For small values of  $\tau_{\text{UO}}$  the numerical precision of these coefficients becomes unreliable (in floating point arithmetic), so for this reason these are replaced with Taylor expansions for  $\tau_{\text{UO}} \lesssim 0.1$ .

An expression for the  $\Lambda^*$  operator can then be devised equivalently to the process used previously for the plane-parallel case. The source function is locally set to unity, and zero elsewhere. This implies that  $c_{\text{U}}$  is also set to 1. Thus the local contribution to the radiation field is given by

$$\Lambda^*(\nu, \vec{d}) = \omega_{c_{\text{U}}} + \omega_{\text{O}}, \quad (6.9)$$

## 6.1 The Formal Solver in Two-Dimensions



**Figure 6.2:** Diagram of sweep order for two-dimensional short-characteristics formal solver.

and remains related to  $\Psi^*$  by the local opacity.

The equivalent integration and  $\Lambda$  operator coefficients for a linear short-characteristics approach can be computed similarly and are analogous to those computed in the one-dimensional case.

### 6.1.2 Evaluation Order and Boundary Conditions

Looking more closely at the order in which the formal solver needs to sweep the grid, we can see that the point O for the ray discussed in the previous section (and shown in Fig. 6.1) would be the 10th node to be solved, and this ordering is shown in Fig. 6.2. Most of the nodes on this figure are solved equivalently to this one, with all necessary quantities known at evaluation time provided the sweep order is preserved, but there are several question marks which require explanation.

The cyan question marks along the  $x$  axis all require values that must be computed from the boundary conditions. The upper and lower boundary conditions in  $z$  are typically taken to be described by a defined in-going radiation field (possibly zero or based on a black body). The upper and lower boundaries in  $x$  can be described by fixed boundary

## 6.1 The Formal Solver in Two-Dimensions

conditions, but it is also common to describe these with periodic boundary conditions where the ray wraps from one side of the grid to the other. The teal question marks along the upper  $x$  boundary can then be interpreted in multiple ways: in the case of periodic boundary conditions they can be prolonged along the teal arrows, and used in the same way as the previously described case. If fixed boundary conditions are used then the intensity at these points must be computed by a linear formal solver indicated by the black arrows ending at the nodes labelled 4 and 8.

Finally, the region around the orange question mark, the tail end of the arrow passing through node 11, requires some additional explanation. In the configuration shown here, the arrow can stop at its intersection with the vertical line on which nodes 6 and 10 lie as the intensity information has been computed at both of these. If an equivalent situation occurs at a periodic  $x$  boundary, then a long-characteristics approach will need to be applied to this ray. This implies that the ray will need to be prolonged back to the previous intersection with a horizontal grid line, where the necessary intensity values can be interpolated, as is shown on this figure. There are multiple options for treating the integration over the segment between this upwind point and the node. Whilst it is possible to take this segment as a singular integration term, very inclined rays may cross multiple vertical grid lines and regions with dramatically varying parameters. For this reason it is common to sub-step along this ray, performing an accumulated short-characteristics integration along each subinterval.

Another problem that may arise, closely related to choosing the correct upwind point, is that of velocity shifts in the medium. van Noort et al. (2002) discussed the possibility of the opacity over an integration integral being underestimated due to differing Doppler shifts at each end of the segment affecting the local opacity by a large margin (say from the core to the wing of the line). As commented by Ibgui et al. (2013), it is also possible for this effect to overestimate the opacity along this segment, in a similar manner. The most common solution to this, proposed by van Noort et al. (2002), and explained in detail in Ibgui et al. (2013) is to subgrid along the ray. The ray is then divided into subintervals along which the velocity may only change by a small (implementation defined) amount relative to the thermal velocity. This approach can be very expensive when large Doppler shifts are present, due to the work involved in computing these segments, interpolating the necessary parameters to each start/end point (this will now require interpolation on both axes, rather than simply one as in the basic case discussed thus far), and the extra numerical integration steps. This subgridding technique is not currently supported in *Lightweaver* but the code was designed with this method in mind, and a sensible location has been left for its implementation.

### 6.1.3 Implementation Details

The restrictive ordering of the formal solver sweep discussed in the previous section imposes constraints on the parallelisation of this algorithm. There is an in-depth discussion of an advanced spatial and frequency parallelisation algorithm for multi-dimensional radiative transfer in Štěpán & Trujillo Bueno (2013), however in *Lightweaver* we assume that the entire simulation domain can be held in memory and the formal solver is parallelised in frequency, equivalently to the plane-parallel case.

The data structures for storing the atmospheric and population information in *Lightweaver* were updated to support two-dimensional atmospheres, storing the data contiguously so as to be able to reuse the core iteration machinery from the plane-parallel case (as inspired by the RH code). Two-dimensional formal solvers can be loaded from external libraries via the same interface as used for their one-dimensional counterparts, and through these interfaces we ensure the modularity of *Lightweaver*. An equivalent interface is also defined for the interpolation function to be used in two-dimensions, giving flexibility in the interpolation order and any form of limiting used (which may need to be adapted to specific grids). *Lightweaver* provides default implementations of the two-dimensional linear and BESSER short-characteristics formal solvers, along with linear and BESSER interpolation schemes for the necessary parameters. The framework defaults to the BESSER formal solver with linear interpolation for the parameters.

For efficiency, the calculation of the ray-grid intersections is performed in a separate pre-pass to the formal solution, as this information can be reused for each formal solution using the same angular quadrature. Whilst it is possible to compute the necessary parameter interpolation weights at this stage, we choose not to do this, as it would render the interpolation interface either more limited, or substantially more complex, due to the need to utilise different numbers of interpolation weights for different schemes. We instead opt to store fractional indices which can be used in conjunction with the grid information in any interpolation procedure. Using 64-bit arithmetic these are a concise and robust method for storing these locations. By design, the intersection calculation is only performed for one upwind and downwind point (excluding long characteristics that cross grid boundaries), as we consider the second order method to be a practical trade-off in terms of computational cost against accuracy. This could easily be updated in the future, and we acknowledge that this is a limitation in terms of the two-dimensional formal solvers that can be loaded via the external interface.

## 6.1.4 Validation

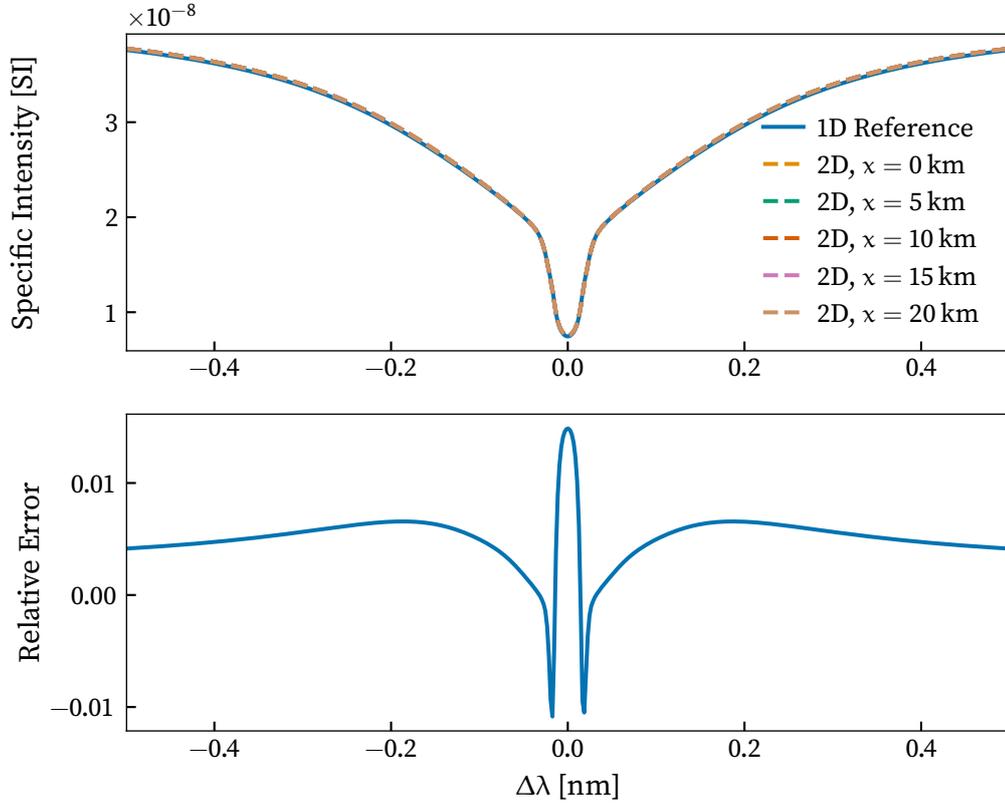
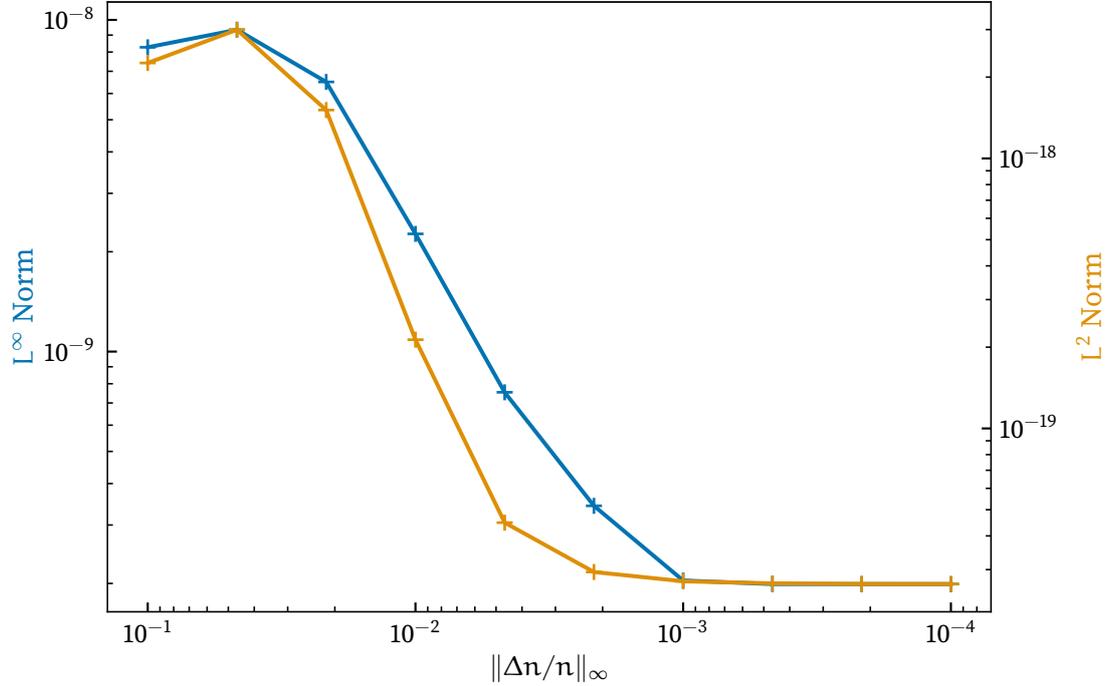


Figure 6.3: Validation of 2D formal solver in static FALC atmosphere with periodic  $x$  boundary conditions.

It is necessary to validate both the two-dimensional formal solver and its integration into the *Lightweaver* framework. Here, we present a basic validation case, a horizontally homogenous FALC atmosphere, using five points in  $x$ , spaced 5 km apart. The boundary conditions in  $x$  are periodic, thermalised at the photosphere, and no radiation is incoming at the top of the atmosphere. The standard configuration of *Lightweaver* in two-dimensions is used, i.e. BESSER formal solver and linear interpolation. The 11th order angular quadrature with 6 rays per octant (Štěpán et al. 2020) is used when solving for the level populations.

In Fig. 6.3 we show the comparison of a Ca II 854.2 nm computed using this two-dimensional method against the same model atom and atmosphere in a plane-parallel configuration. Both of these are iterated until the maximum relative change in the Ca II populations

## 6.1 The Formal Solver in Two-Dimensions

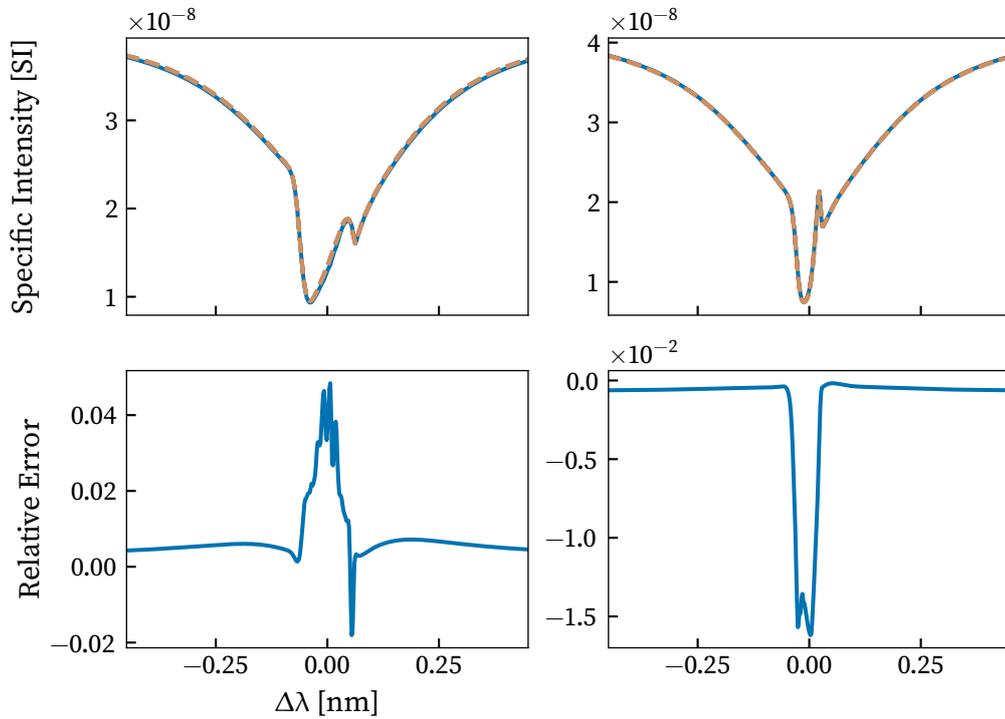


**Figure 6.4:**  $L^\infty$  and  $L^2$  norms of the difference of the Ca II 854.2 nm line profiles between the 1 and 2D formal solvers (including population iteration procedure).

is less than  $10^{-4}$ . The outgoing radiation shown in the upper panel of this figure was synthesised using rays in the  $x - z$  plane with  $\mu_z = 0.9$ . Visually, the dashed lines, showing the solution from different  $x$  locations using the two-dimensional formal solver overlie each other perfectly, showing that the output is homogenous along the  $x$  axis, as is to be expected from this horizontally homogenous atmosphere with periodic boundary conditions. There is a slight visible offset between the solutions generated from the plane-parallel and two-dimensional simulations, and the relative difference between these is shown in the lower panel of this figure. We see that the relative difference is greatest in the line core, peaking at around 1.5%. Overall, this is very good agreement considering the different underlying methods used in the formal solvers.

In Fig. 6.4 we show how the difference between the plane-parallel and two-dimensional models changes as the  $L^\infty$  norm of the relative population change decreases (i.e. as we approach convergence). Further iterations beyond a maximum relative population change of  $10^{-3}$  do not substantially improve the agreement between the two methods. Thus, the situation shown in Fig. 6.3 is sufficiently converged to compare the final solutions of the

## 6.1 The Formal Solver in Two-Dimensions



**Figure 6.5:** Validation of 2D formal solver in FALC atmosphere with  $20 \text{ km s}^{-1}$  sinusoidal vertical velocity field and periodic  $x$  boundary conditions in the left-hand column and  $5 \text{ km s}^{-1}$  vertical sinusoidal velocity field and fixed  $x$  boundary conditions computed from a one-dimensional plane-parallel model in the right-hand column. The left-hand column is again synthesised at  $\mu_z = 0.9$  whereas the right-hand column is synthesised at  $\mu_z = 0.99999$  to include transverse effects, without simply sampling the fixed boundary condition that would give the plane-parallel result directly.

two methods.

We find that the two-dimensional formal solver performs well, both in terms of accuracy and computational performance. For the example presented here the plane-parallel solution takes 1.026 s and the two-dimensional solution takes 12.471 s after 55 and 266 iterations respectively (this timing includes configuring the additional contexts and the formal solutions used for the convergence analysis shown in Fig. 6.4). Both of these simulations were run with 16 threads on *tomahna*.

Similar validation tests have been run for fixed boundary conditions and Doppler shifted atmospheres, all of which yielded extremely satisfactory results. These are shown in Fig. 6.5 and utilise the same configuration other than the parameters discussed below. The

## 6.1 The Formal Solver in Two-Dimensions

left-hand column shows the effects of a vertical velocity field on a simulation with periodic boundary conditions. This velocity field is uniform in  $x$  and defined by the additive inverse of a single period sine wave spanning the entire FALC atmosphere, with the first and last three points set to  $0 \text{ km s}^{-1}$ . The amplitude of this wave is set to  $20 \text{ km s}^{-1}$ . As such, it represents a simpler version of the plane-parallel test presented in Fig. 4.2. The lower row of this column shows that there is once again good agreement between the plane-parallel and two-dimensional models, with differences peaking around 5% in the line core, but the line shape being accurately reproduced. This difference could likely be reduced by the use of the previously discussed subgridding technique, but as the simulations that are presented later in this chapter do not use velocity fields within the 2D slab, this work was not undertaken.

In the right-hand column a model with fixed  $x$  boundary conditions is presented. This model also applies an equivalent velocity field, albeit with a smaller  $5 \text{ km s}^{-1}$  amplitude. The boundary conditions are computed from the associated one-dimensional simulation; the radiation along each ray of the two-dimensional quadrature needed for the boundary condition at each depth used in the two-dimensional simulation is synthesised from the plane-parallel simulation for both boundary conditions along the  $x$  axis. The radiation presented here is synthesised along a ray with  $\mu_z = 0.99999$  ( $0.25^\circ$ ) so that the ray through central  $x$  cell at the top of the model does not intersect directly with either of these boundary conditions (as this would simply be sampling the plane-parallel simulation), whilst still testing the two-dimensional nature of the formal solver. The agreement here is once again very good, with the error peaking around 1.5%.

The coupling between the two-dimensional and plane-parallel models utilised here is facilitated by the design of the *Lightweaver* framework, and will be leveraged extensively in the following sections. A *Lightweaver* model describing a plane-parallel boundary condition can be updated and synthesise its output dynamically in response to the rays requested from the two-dimensional model, without the need to precompute these values and go through a multi-step saving and loading process that is prone to error when simulation parameters change. This can be especially difficult in cases where the incoming radiation field from a fixed boundary condition is allowed to be anisotropic, requiring agreement between the rays in both models. The boundary condition radiation for the almost vertical rays used in the right-hand column is computed automatically due to the design of the class that describes the boundary condition, and the coupling allowed by the framework, which allows easy manipulation and coupling of multiple radiative transfer contexts within the same program.

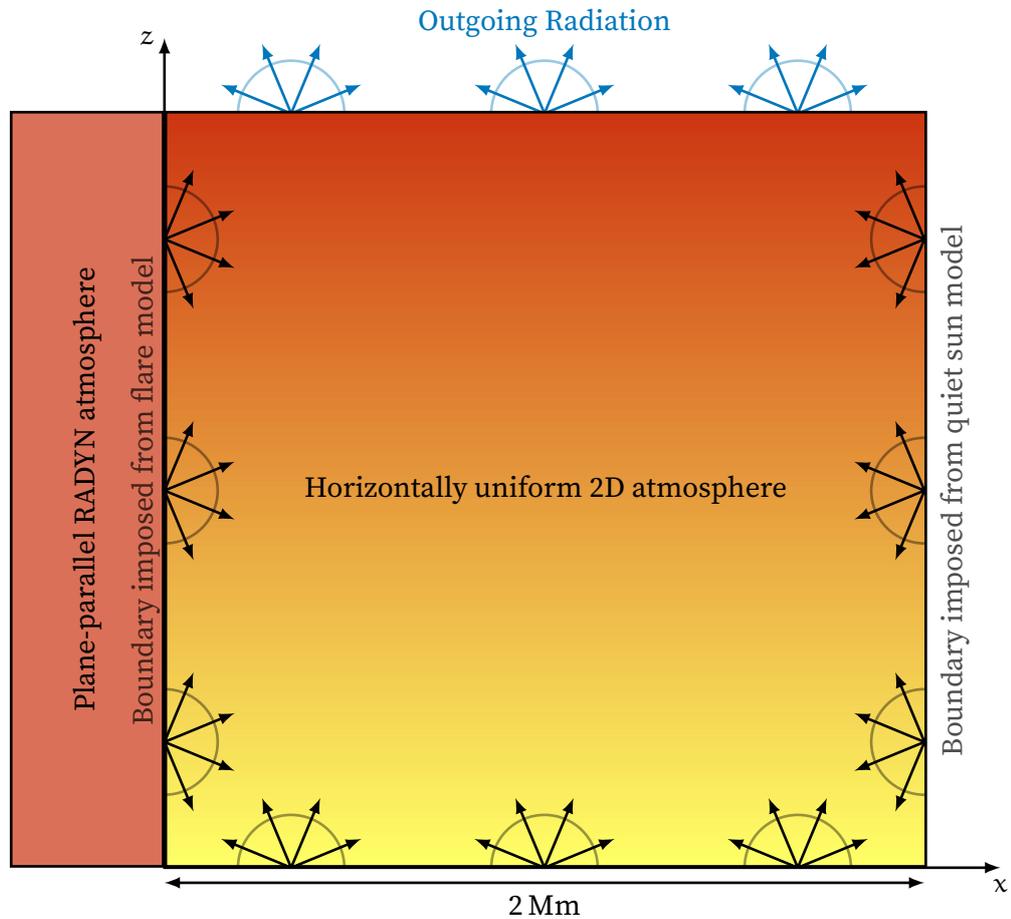
## 6.2 2D Simulation Configuration

Flares produce huge changes in the radiation field, in both lines and continua. They are typically modelled in a plane-parallel context and we analyse the radiation leaving the top of this plane-parallel atmosphere. In reality, the flaring kernels that we are simulating with these RHD models are likely small; on the scale of 10s to 100s of km in diameter (Jing et al. 2016) and represent an estimate of the conditions in the core of a heated flux-tube. As the resolution of solar telescopes increases, so does their ability to resolve spatial effects tangential to the solar surface (henceforth horizontal). The lack of horizontal atmospheric homogeneity, that is not accounted for in these models, may produce complex intensity structures resolvable with these new telescopes as the huge outgoing flux of radiation from the flare core impinges on and interacts with neighbouring plasma. These effects include the possible ‘core-halo’ pattern reported around flare kernels in ground-based line and continuum imaging (Neidig et al. 1993; Xu et al. 2006), TRACE white light observations (Hudson et al. 2006), and Hinode/Solar Optical Telescope (SOT) G-band and Fe I 630.2 nm observations (Isobe et al. 2007). The proposed explanation for these halos is typically radiative backwarming from the flare’s radiation (previously investigated in-depth using Mg I lines by Metcalf et al. (1990b)). In the following we shall focus primarily on NLTE radiative effects affecting plasma neighbouring a flare, but leave the inclusion of temperature variations due to absorption of radiation to a future study.

It is reasonable to suggest that the plasma neighbouring the flare is substantially cooler, both due to the lack of direct heating, and the low strength of cross-field conduction relative to that along the field (Spitzer & Härm 1953), however photons are not affected by these limitations. Thus, we will investigate the effects of illumination from a neighbouring plane-parallel flare model on a two-dimensional slab of plasma representing the quiet atmosphere, with a time-dependent radiative treatment of both of these. This model can then serve as a “first-order” approximate investigation of the effects of flare radiation on the slab, as well as the depth of radiation penetration, effects on the atomic populations, and observable signatures.

Modelling undertaken by Leenaarts et al. (2012a) suggests that a three-dimensional treatment of H $\alpha$  in quiet Sun models yields very different synthetic line profiles than a 1.5D column-by-column plane-parallel treatment. Bjørgen et al. (2019) found that these differences were less significant for the bright structure of an active region (based on the simulations of Cheung et al. 2019), but definition appears to be lost in regions adjacent to this structure when a 1.5D treatment is applied. It is precisely the effects of a bright region

## 6.2 2D Simulation Configuration



**Figure 6.6:** Configuration of the two-dimensional simulation showing the flaring boundary condition.

on the neighbouring atmosphere that we wish to investigate, and a simple two-dimensional model should reveal the importance and possible time-dependence of reactions to this radiation.

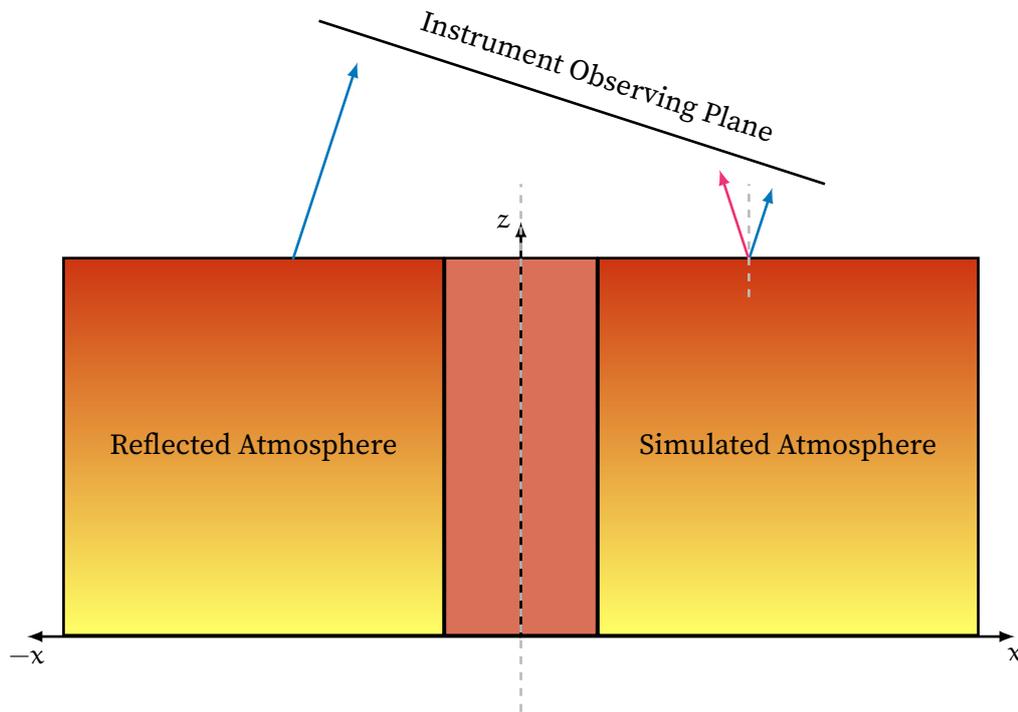
Our simulation is set up as shown in Fig. 6.6: the primary simulation domain is a 2 Mm wide slab of plasma initially set to the quiet sun atmosphere used for the RADYN simulation. On one side of this slab we place the RADYN simulation, and compute the intensity along each ray of the angular quadrature used for the 2D slab, at each depth in the simulation. The other  $x$  boundary is treated equivalently, but using the fixed initial quiet sun atmosphere from the RADYN simulation and the 2D slab. The width of this domain was chosen by a manual iterative process ensuring that the line profiles emerging from the slab close to the quiet Sun boundary naturally and smoothly return to close to the quiet Sun values, otherwise this fixed boundary will be spuriously sinking large quantities of energy.

## 6.2 2D Simulation Configuration

Similarly to the process described in the previous chapter a time-dependent simulation is run, again reprocessing the thermodynamic atmospheric properties using RADYN's internal timestep. The separate components of this model share a  $z$  stratification based on a combination of RADYN's grids used for both the initial quiet sun atmosphere and the current timestep. This method ensures that 450 points are spaced across the entire altitude range of the atmosphere and provide sufficient resolution for the transition region of both the quiet sun model *and* that of the flare model. The populations determined by *Lightweaver* are interpolated between the  $z$  grids from one timestep to the next, and locally scaled to follow the mass density (this is typically a small adjustment, but without it errors can grow as points move through the transition region). The electron density in the flare model is loaded from the RADYN output, and charge is conserved in the 2D slab using the secondary Newton-Raphson iteration procedure discussed in Sec. 2.4.13. The 6 rays per octant of the unit sphere quadrature of Štěpán et al. (2020) was chosen, as the plasma in the 2D slab is static and these rays capture enough detail to describe the radiation field leaving the plane parallel model (where the radiative transfer model natively uses 5 Gauss-Legendre rays per quadrant of the unit disc). In addition to the 450 points in  $z$ , we use 41 linearly spaced points in  $x$  to discretise the two-dimensional atmosphere. As these points span 2 Mm each point is spaced 50 km apart. The use of this grid is justified in Sec. 6.3.5. Due to the very fine  $z$  spacing that often occurs due to the strong gradients in the transition region, many of the 2D cells have an aspect ratio very far from square, which can pose convergence difficulties if insufficient angular resolution is used. Unlike the model in the previous chapter, we do not consider advection here: the plasma in the two-dimensional slab is static, and the effect of advection on the flaring boundary condition is small. This also has the effect of allowing more flexibility in the  $z$  stratification, as a grid that is stable for radiative transfer may not be so for advection.

The left-most column of the 2D slab requires special treatment due to the nature of fixed boundary conditions in radiative transfer simulations. The incoming radiation from the flare model is specified for each ray and depth in the first column of the atmosphere. As the intensity is specified for all incoming (rightgoing) rays here, it cannot be calculated taking into account the local parameters, meanwhile the outgoing (leftgoing) rays *do* take into account the local parameters, and then inform the local operator acting on these populations. Thus, if this column has its thermodynamic properties fixed to those of the initial quiet sun model, it behaves as if it is only receiving radiation from the right, whilst only being affected by the flare in a “second-hand” sense. This leads to a dark first column in the two-dimensional synthesis. To minimise these effects we copy the flare atmosphere and populations into this first column, and hold these fixed over each timestep. These

## 6.2 2D Simulation Configuration



**Figure 6.7:** Using the single simulation shown in Fig. 6.6 to investigate the radiation observed by a slit spectrograph looking across the flaring region. The radiation emitted along the blue arrow from the reflected atmosphere is the same as that from the magenta arrow in the simulated atmosphere. The angles of the rays shown in this diagram are exaggerated relative to those used in the results presented later, to better illustrate the configuration.

populations are then consistent with the adjacent plane-parallel atmosphere and the radiation emerging from it which is used as a fixed boundary condition. There should not be any need to perform the same process at the quiet sun boundary, as it is placed 2 Mm away and should change very little over the course of the simulation, remaining consistent on both sides of the boundary.

This configuration produces the outgoing intensity at 12 different angles from each of the 41 cells in  $x$ . These rays are symmetric in the  $z$ -axis, so we can use the symmetric definition of these to synthesise the radiation observed by a theoretical slit spectrograph viewing the sun from a particular inclination with the flare in the centre of the slit. This is shown in Fig. 6.7 for an observation inclined as shown we are observing the radiation along the blue arrows. Reflecting the entire simulation in the  $z$ -axis, the intensity observed along the blue ray from the reflected atmosphere is the same as the radiation along the magenta ray produced in the original simulation configuration shown in Fig. 6.6. Using the information produced by the original simulation configuration we can therefore produce this view of

### 6.3 Simulation Results

both sides of the flare.

In the following results, we consider that the flaring boundary is at  $x = 0$  and will be presented as infinitesimally narrow, with the simulated atmosphere in positive  $x$ , and its reflection in negative  $x$ . To better describe the atmosphere we will split it into regions based on  $x$ -coordinate as follows:  $A := (0, 0.5]$  Mm,  $B := (0.5, 1]$  Mm,  $C := (1, 2]$  Mm. We will use minuscules (lower-case) of these designations to denote the reflections of their associated regions in the  $z$ -axis (i.e. in negative  $x$ ).

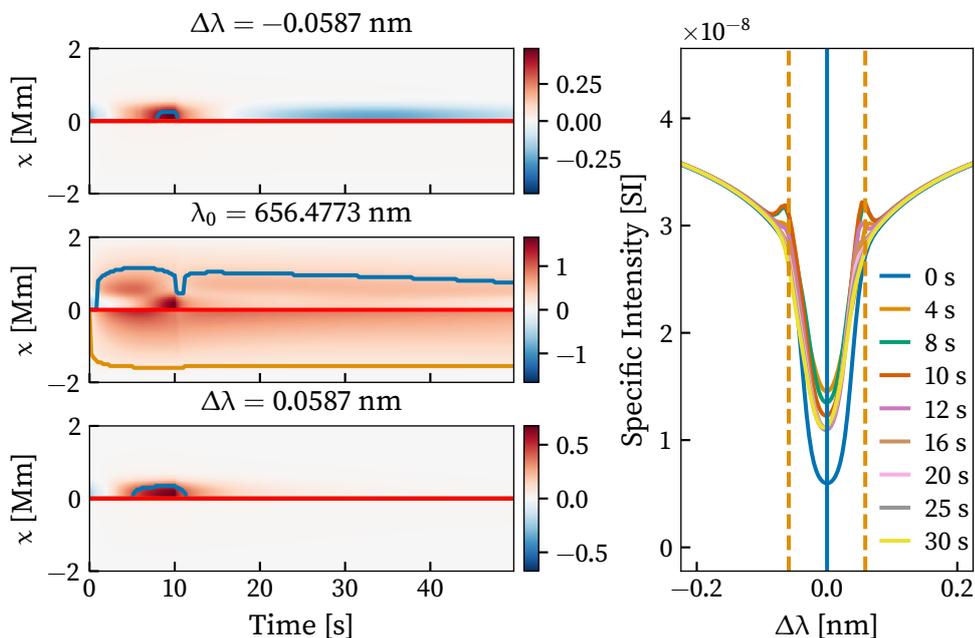
We will focus primarily on these inclined observations as flares are extremely unlikely to occur at exactly disc centre and the effects of inclination are therefore important. Even a slight inclination can have a large effect, and in the following we will focus primarily on the most vertical ray present in our quadrature, with an angle of approximately  $18^\circ$  ( $\mu_z \approx 0.95$ ) to the surface normal. Only  $\sim 1/25$  of the visible solar disc has a viewing angle smaller than this, so inclination effects will be at least this significant for the majority of observed flares. The two-dimensional model is assumed to be homogenous along  $y$ , and the rays of the angular quadrature used for integration have non-zero  $y$  components to correctly sample the unit sphere. The most vertical ray has an inclination from the surface normal of  $18^\circ$  and an azimuth from the  $x$ -axis of  $83^\circ$ , meaning its projected inclination in the non-homogeneous  $x - z$  plane is  $2.24^\circ$ .

In this simulation we use the same model atoms as the previous chapter, with both hydrogen and calcium being set as active species. This simulation is carried out for two different flare models illuminating the 2D slab: F9 and F10 models with identical heating parameters to those of Chapter 5 including the Lyman line photoionisation effects, but with RADYN's in-built coronal irradiation enabled. We will now show the results of these simulations, looking at both the observable effects and the population changes internal to the slab, focusing on the  $H\alpha$  and Ca II 854.2 nm spectral lines, before comparing these to observations taken with the CRISP instrument on the SST. To justify the importance of the time-dependent treatment, we will also compare these results against statistical equilibrium solutions computed at several timesteps.

### 6.3 Simulation Results

Due to the additional complexity of the two-dimensional simulation these models are substantially more computationally intensive than the plane-parallel models of the previous chapter, both due to the larger number of points at which the RTE is to be solved as well as

### 6.3 Simulation Results



**Figure 6.8:**  $H\alpha$  for 2D F9 case. The left-hand column shows the intensity enhancement over quiet values on each side of the flare (red line) for the three different wavelengths indicated on the right-hand panel with vertical lines. The blue lines show the extent of a 40% enhancement over the quiet boundary condition, and the orange a 10% enhancement. The right-hand panel shows the spectral output at position  $x = 500$  km at different times in the simulation.

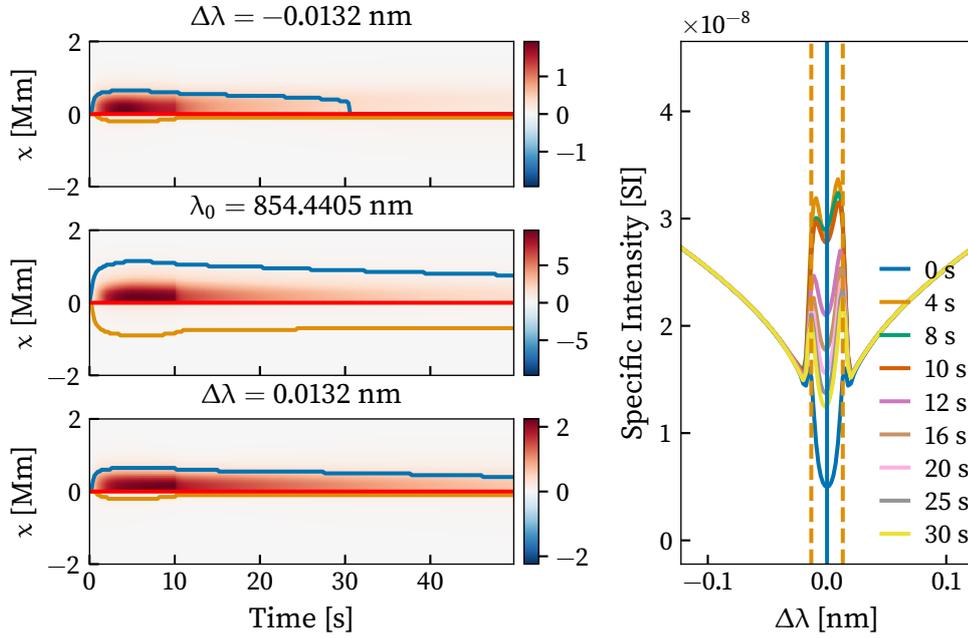
the increased angular samples and interpolations needed at each point. The F9 simulation, consisting of 1793 internal timesteps takes  $\sim 2,300$  CPU hours, and the F10 model, with 5996 timesteps takes  $\sim 8,000$  CPU hours on *hercules*. The tool employed for these simulations is derived from the code described in Chap. 5, and is available on GitHub<sup>2</sup>, with archival on Zenodo (Osborne 2021c).

#### 6.3.1 Observed Radiation

Figs. 6.8–6.11 show the  $H\alpha$  and Ca II 854.2 nm spectral lines in the F9 and F10 simulations respectively. The left-hand column shows the intensity enhancement in three different wavelengths as a function of space and time, configured as described in Fig. 6.7 with the reflected side shown in negative  $x$ . The flare model is presumed to have an infinitesimal width (in  $x$ ) indicated by the horizontal red line. These three wavelengths sample the line

<sup>2</sup><https://github.com/Goobley/MsLightweaver2d>

### 6.3 Simulation Results

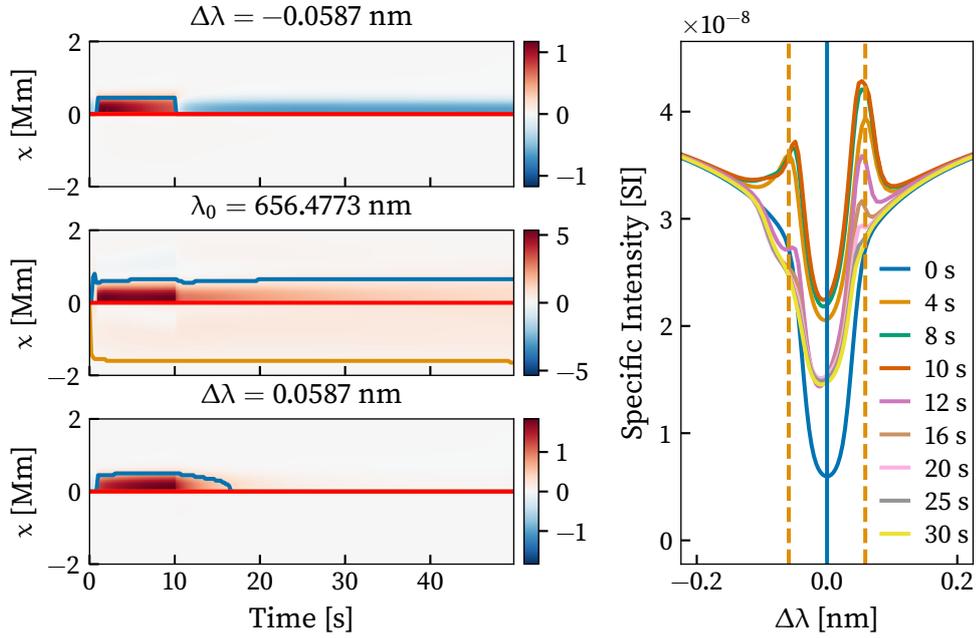


**Figure 6.9:** Ca II 854.2 nm for 2D F9 case. The panels show equivalent information to Fig. 6.8.

core and secondary features on either side of the line and are indicated by the vertical lines in the left-hand panel, with the line core indicated in blue and the two secondary wavelengths in dashed orange. This left-hand panel also shows the emergent intensity in this line at  $x = 500$  km from the flare at different times throughout the simulation for the most vertical emergent ray from the simulation. Blue and orange lines are also plotted on the left-hand panel and indicate the distance at which an enhancement of 40% and 10% (respectively) over the quiet boundary are observed for this wavelength and time in the simulation. The different thresholds are used for the positive and negative  $x$  regions respectively as the scale of the enhancements vary significantly between these, with the effect in the negative  $x$  region being notably smaller.

It is clear from these figures that in all cases the enhancement in the line core is significantly larger in all of value, extent (dimension in  $x$ ), and duration than the wing feature enhancements, with both the positive and negative  $x$  enhancements remaining over 0.5 Mm from the flare for the entire duration. The enhancements in the wing features are much smaller in extent and more transient, especially for H $\alpha$ , where the blue wing feature near the flaring boundary enters a reduction relative to the quiet boundary value soon after heating ends in the flare model ( $t = 10$  s). In Ca II 854.2 nm both the blue and red wing features in positive  $x$  present a similar extent, but the red wing feature persists

### 6.3 Simulation Results



**Figure 6.10:** H $\alpha$  for 2D F10 case. The panels show equivalent information to Fig. 6.8.

for the entire simulation whereas the blue feature vanishes around  $t = 30$  s. There are interesting differences in the negative  $x$  enhancements for Ca II 854.2 nm between the F9 and F10 models. In the former there is an enhancement only during the flare heating period, whereas in the latter there is no significant enhancement until after the heating has ended, which then persists for the remainder of the simulation.

The H $\alpha$  line profile in the F9 simulation (Fig. 6.8) synthesised at  $x = 500$  km shows that over the course of the simulation the line core intensity increases, but does not rise above the continuum value, and two secondary features form in the red and blue wings during the flare heating. These decay rapidly after heating ends, dropping to below the initial quiet value at our blue wing sampling. The Ca II 854.2 nm computed in this same simulation (Fig. 6.9) emits strongly as a doubly-peaked line with a slight asymmetry in favour the red wing and remains so for the majority of the simulation, decaying slowly over time, with the line core decaying more rapidly than the secondary peaks (the peaks of the line profile in this centrally-reversed state), leading to an increasingly deep central reversal.

The effects are similar, but much larger in the F10 simulation. The H $\alpha$  line (Fig. 6.10) is greatly enhanced presenting a doubly-peaked shape with a deep central reversal during heating. This line is notably asymmetric in favour of the red wing throughout the

### 6.3 Simulation Results

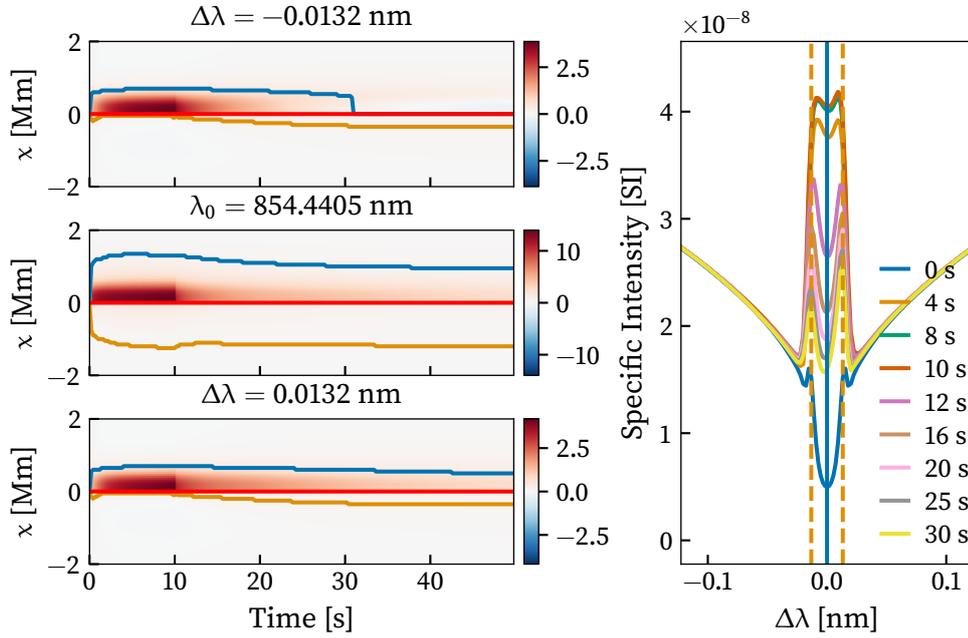


Figure 6.11: Ca II 854.2 nm for 2D F10 case. The panels show equivalent information to Fig. 6.8.

simulation. After the flare heating ends, the red secondary peak remains enhanced and decays slowly whereas the blue wing rapidly drops below the initial quiet value. The Ca II 854.2 nm line (Fig. 6.11) is a more dramatically enhanced version of the one considered previously for the F9 simulation although its asymmetry flips in favour of the blue wing for the  $t = 12$  s line profile.

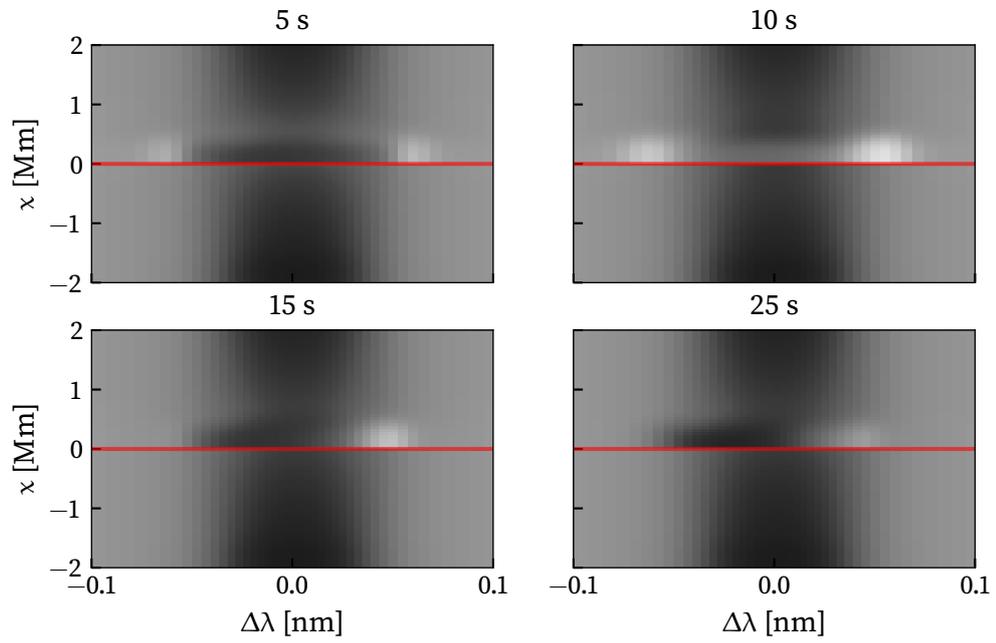
We note that for both spectral lines the far wings and continuum ( $|\Delta\lambda| \geq 0.1$  nm for H $\alpha$  and  $|\Delta\lambda| \geq 0.04$  nm for Ca II 854.2 nm) have very small or no variations throughout both simulations. This will be discussed later, when the effects on the populations are investigated.

To better interpret these simulations we can also look at what our theoretical slit spectrograph described in Fig. 6.7 would observe at different times in the simulation.

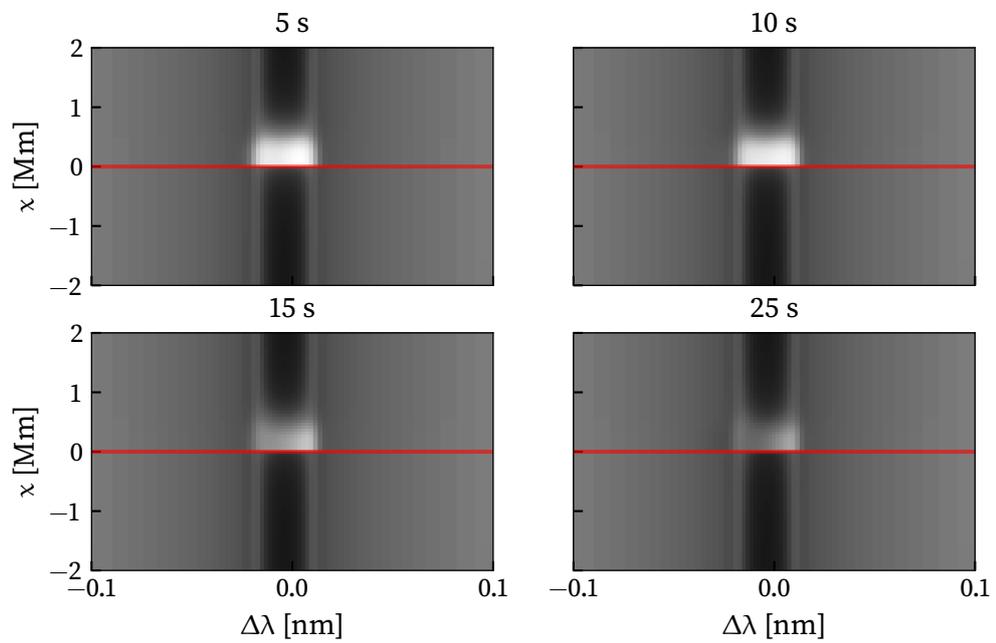
#### 6.3.2 Spectroscopic Results

The emergent spectra from the F9 simulation are shown for H $\alpha$  and Ca II 854.2 nm in Figs. 6.12 and 6.13 respectively. The intensity scale is constant for all timesteps for each line. As expected from the previous imaging results, significant variations in the line profiles

### 6.3 Simulation Results



**Figure 6.12:** H $\alpha$  intensity on both sides of the flare (indicated by the red line) at different points during the F9 simulation. The scale of the colourmap is constant across the panels.



**Figure 6.13:** Ca II 854.2 nm intensity in the F9 simulation shown equivalently to Fig. 6.12.

### 6.3 Simulation Results

are observed 0.5–1 Mm from the flare location. Once again, there is a large difference between the positive and negative  $x$  regions, with much greater variations observed in the former. This is likely due to rays passing through the positive  $x$  region encountering more scattered light from the flare and directly intersecting the flaring atmosphere. There is some enhancement in the negative  $x$  region, presenting as a narrowing of the line core in  $H\alpha$  and enhancement of the core of Ca II 854.2 nm, but it is much less significant than that present in the positive  $x$  region. Despite the horizontal uniformity of plasma properties in this atmosphere, there is substantial variation in the observed radiation.

Starting with the  $H\alpha$  line profiles shown in Fig. 6.12, during heating ( $t = 5$  s) we see a line profile in region A with a broadened core in absorption (dark in the figure) and enhanced secondary peaks in both the red and blue around  $|\Delta\lambda| = 0.06$  nm. In B, the line core is enhanced relative to both A and its quiet value with the secondary peaks in the wings around decaying rapidly, producing a shallower absorption profile, before a smooth return to the quiet sun line profile over C. In the negative  $x$  region the line is less deep near the flare, decaying back to pre-flare levels over the 2 Mm extent. The situation is similar at the end of the heating ( $t = 10$  s), but here the core is more significantly enhanced in region A before deepening but remaining well above the quiet value over B and smoothly transitioning back to its quiet sun form in region C. A slight asymmetry of the red wing (due to the brightening around  $\Delta\lambda = 0.05$  nm) starts to appear here in region A. The negative  $x$  region remains similar to  $t = 5$  s although the line core is less enhanced immediately adjacent to the flare. Some time after the heating has ended ( $t = 15$  s and  $t = 25$  s), the previous brightening (secondary peak) in the blue wing over A is now a dimming, but there remains an enhancement in the red wing, which fades gradually. In region B the line appears shallower than for  $t = 10$  s, but not as enhanced as for  $t = 5$  s. The negative  $x$  region and C remain similar to the  $t = 10$  s line profile, with the line core smoothly increasing in depth until reaching the quiet sun value as distance from the flare increases.

The Ca II 854.2 nm line profiles in the F9 simulation shown in Fig. 6.13 tell a similar story, although the line width does not vary noticeably throughout the simulation. The line is strongly enhanced in emission over A for  $t = 5$  s and  $t = 10$  s, with a slight red asymmetry. At later times this line remains in emission with a more significant red asymmetry. Region B presents the start of a smooth transition back towards the quiet sun line profile, but retains asymmetry present in region A. The negative  $x$  region presents slight core enhancement over a, which is stronger during heating, before decaying slowly after heating. Regions b, c, and C do not change substantially over the course of the simulation.

### 6.3 Simulation Results

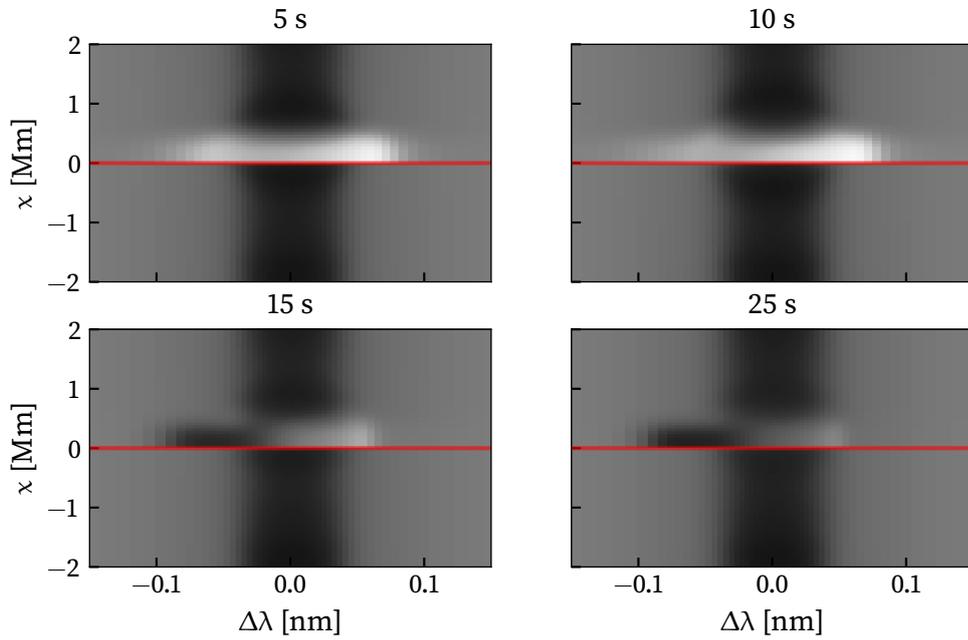


Figure 6.14: H $\alpha$  intensity in the F10 simulation shown equivalently to Fig. 6.12.

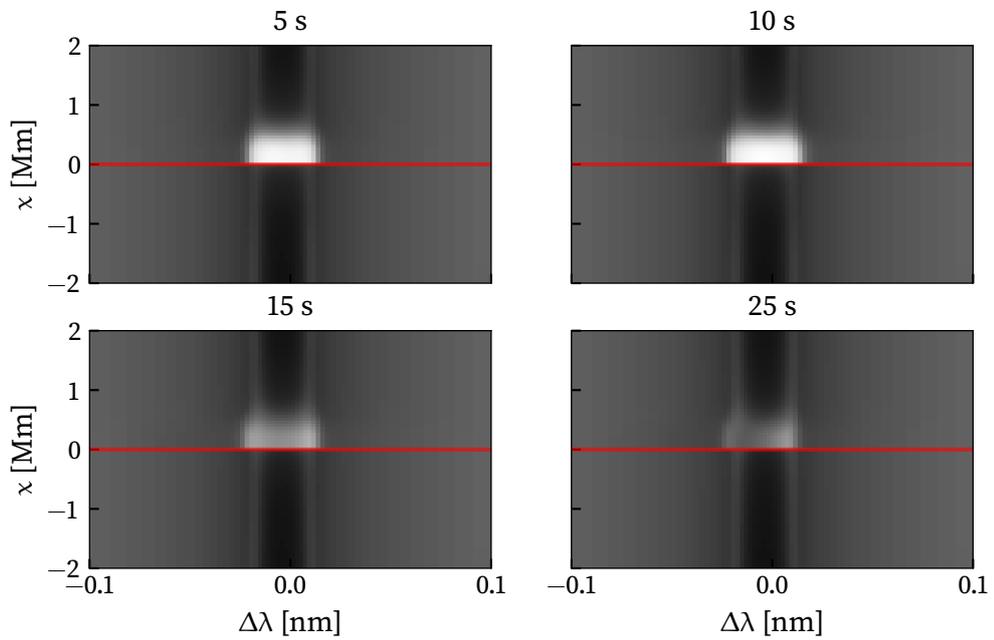


Figure 6.15: Ca II 854.2 nm intensity in the F10 simulation shown equivalently to Fig. 6.12.

### 6.3 Simulation Results

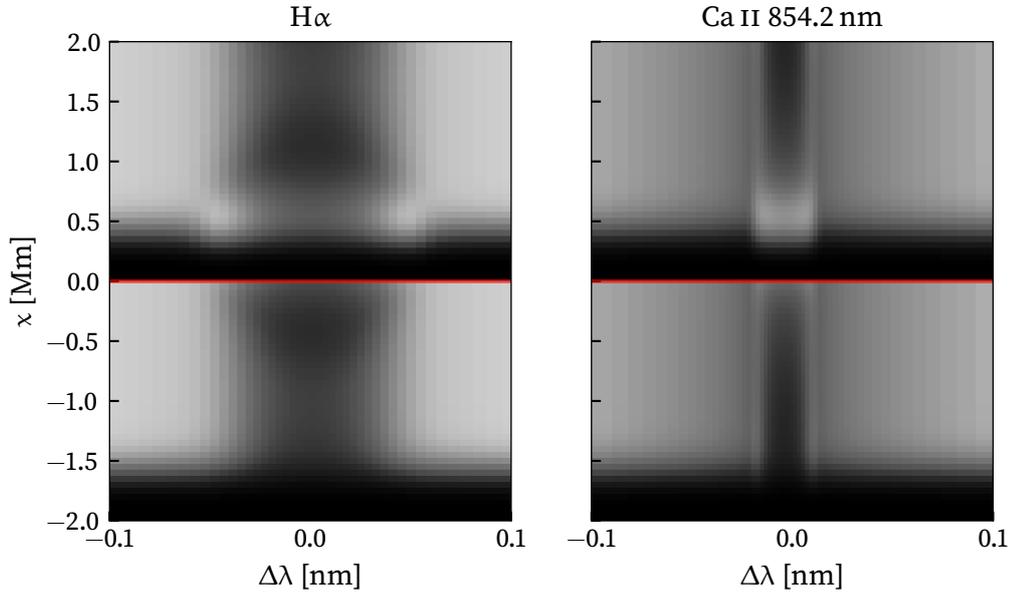
The  $H\alpha$  and Ca II 854.2 nm line profiles from the F10 simulation are shown in Figs. 6.14 and 6.15 respectively. The results are similar to those from the F9 simulations, but mostly exaggerated. The line profiles over A and B are substantially broadened relative to those shown in the F9 simulation, which is not surprising, as much of this radiation comes from the flaring boundary condition, which is heated by a significantly more energetic beam. During the cooling phase there is still a strong red asymmetry in this region; much stronger than that in the F9 model. This is likely due to the higher velocities present in the flaring boundary. This brightening encompasses the red wing, line core and enters the blue wing in region A. Focussing on  $H\alpha$  (Fig. 6.14), the broadening in B, which is much more significant than that seen in the F9 case, is not constant in  $x$ . For  $t = 5$  and 10 s the line profile is enhanced over quiet values and substantially broadened, decaying slowly back to quiet values as the distance from the flare increases, whereas for  $t = 15$  and 25 s, this effect increases over the first half of B before reducing over the second half. This reduction continues into C with the line appearing at its narrowest around  $x = 1.5$  Mm, although this feature's distance from the flare decreases to  $\sim 1.2$  Mm at  $t = 25$  s.

This broadening is somewhat mirrored in the negative  $x$  region. At early times the line appears significantly broadened over a, before decaying back towards the quiet sun profile over b and c. The variation over b and c is most significant for  $t = 10$  and 15 s but remains important at  $t = 25$  s.

Similarly to the F9 case, the Ca II 854.2 nm profile is most affected over A, presenting substantial broadening, initially symmetric but with a red brightening ( $\Delta\lambda \approx 0.03$  nm) appearing during the cooling phase. During this phase the line remains broadened in the blue wing whilst the red wing half-width returns to a value similar to that found in quiet sun profile, despite the intensity enhancement. These broadenings persist into B and rapidly, but smoothly, decrease to the constant line width observed in the F9 case. The variations in the intensity of the line continue into C and decay smoothly back towards the quiet sun line profile. The negative  $x$  region is also similar to the F9 case, with the more significant effects appearing in region a. The line core near the flare is shallower and narrower than in the F9 case, but this once again smoothly returns to values close to the quiet sun line profile over the course of a, with only a slight enhancement in the line core persisting into b.

We see that these optical line profiles can present significant core enhancements up to 1 Mm away from the flare boundary in these simulations when looking at a viewing angle that intersects this boundary (i.e. the positive  $x$  region). For viewing angles where this is not the case (the negative  $x$  region), the effects on the line profile are much more modest,

### 6.3 Simulation Results



**Figure 6.16:** Intensity at  $t = 10$  s in the F10 model for both  $H\alpha$  and Ca II 854.2 nm with the boundary conditions in  $x$  set to emit no radiation.

but measurable line core enhancements can still be detected over 1 Mm from the flare with wing effects up to 0.4 Mm from the flaring boundary.

As some of the rays emerging from positive  $x$  region intersect the flaring boundary condition it is important to ascertain to what extent we are observing the “conventional” plane-parallel simulation, and what is due to two-dimensional effects. To assess this we have synthesised the spectra of the  $H\alpha$  and Ca II 854.2 nm lines using the populations computed for  $t = 10$  s in the F10 simulation, but setting the radiation from the boundary conditions to zero. These spectra are shown in Fig. 6.16. For both lines, in region A, there is no emission in the far wings and continuum ( $|\Delta\lambda| \geq 0.07$  nm for  $H\alpha$  and  $|\Delta\lambda| \geq 0.02$  nm for Ca II 854.2 nm) below 0.25 Mm. For  $H\alpha$ , the secondary peaks in the wings ( $|\Delta\lambda| \approx 0.06$  nm) protrude approximately 100 km further toward the flaring boundary with the line core protruding approximately 50 km less than these secondary peaks. The situation is reversed for Ca II 854.2 nm: the line core protrudes similarly to the secondary  $H\alpha$  peaks, and there is a slight recess of approximately 50 km at wavelengths just beyond the secondary peaks ( $|\Delta\lambda| > 0.02$  nm) which returns smoothly and rapidly towards the continuum values. There is no significant variation in extent in the negative  $x$  region, with visible emission starting around 0.3 Mm from the quiet boundary.

These distances are quite consistent with an approximate calculation using the ray dir-

### 6.3 Simulation Results

ection. The ray's projected inclination from the  $z$  axis in the  $x - z$  plane is  $2.24^\circ$  and the simulation has a vertical extent (from RADYN) of approximately 10 Mm. The transition region occurs at approximately  $z = 1.75$  Mm. So a ray entering the top boundary will have a horizontal displacement of approximately 0.3 Mm before entering the upper chromosphere, where the plasma starts to become optically thick at these wavelengths.

There are clear effects due to the two-dimensional simulation beyond 0.25 Mm from the flaring boundary, and these lie well within the regions of important variation outlined in our imaging and spectroscopic analyses of the data. To further investigate the flare's effect on the slab and why the continuum emission is not affected we will now see how the atomic populations within the slab vary.

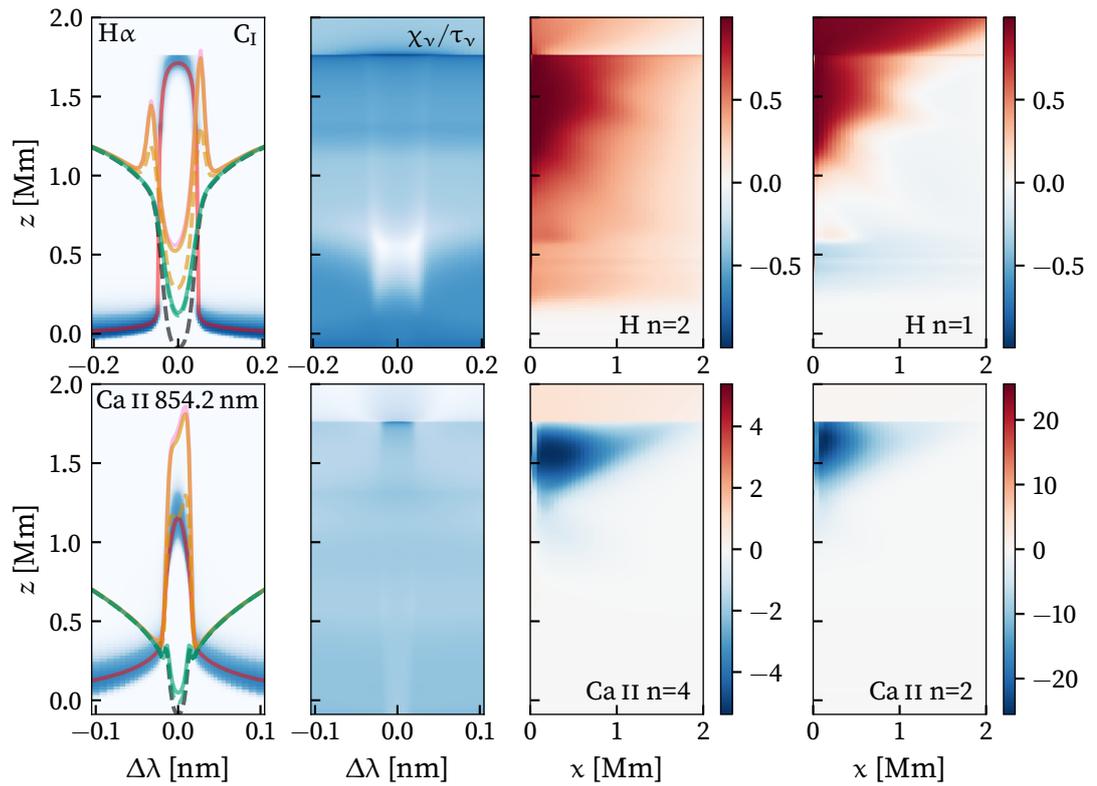
#### 6.3.3 Population Changes

To investigate how and where the flare's radiation field affects the plasma in our two-dimensional slab we can compare the atomic populations responsible for these spectral lines against their initial statistical equilibrium values. For the F9 simulation Fig. 6.17 shows the contribution function  $C_I$  in the first column, the effect of local opacity on the outgoing radiation ( $\chi_\nu/\tau_\nu$ ) in the second column, and the relative change of the upper and lower levels of the transitions considered from their initial statistical equilibrium values in the last two columns respectively. The relative change of the population is computed as

$$\frac{n_{i,t} - n_{i,0}}{n_{i,t}}, \quad (6.10)$$

where  $n_{i,t}$  is the number density of atomic level  $i$  at time  $t$ . The indexing of the levels shown in the plots is the indexing used by *Lightweaver's* model atoms, which is zero indexed. These properties are all shown at  $t = 10$  s. The contribution function and local opacity effect are computed by treating the atmosphere close to the flaring boundary at  $x = 50$  km as plane-parallel and computing the contribution function in the traditional way. This approach was taken due to the difficulty formulating the contribution function in the two-dimensional slab: the wavelength varying contributions of even a single output spectrum from a two-dimensional atmosphere cannot be represented directly on a two-dimensional plot, and this problem of dimensionality becomes greater still when investigating the contribution to each outgoing pixel. We therefore use this simplified model as it should still inform us of the primary line forming and absorbing regions, especially as the thermodynamic structure of the two-dimensional slab is constant (other than the first column, copied from the flaring boundary). The  $\chi_\nu/\tau_\nu$  plot is computed equivalently to the contribution

### 6.3 Simulation Results



**Figure 6.17:** Contribution functions,  $\chi_v/\tau_v$ , and relative changes to the upper and lower level populations from initial values for the H $\alpha$  and Ca II 854.2 nm lines in the F9 simulation at  $t = 10$  s. The  $\tau = 1$  line and initial line profile are plotted in red and dashed black respectively in the first column. The pink line profile shows the flaring boundary condition. The other solid and dashed line profiles in the first panel are synthesised at 200, and 400 km from the flaring boundary respectively, with orange being positive  $x$  and green being negative  $x$ .

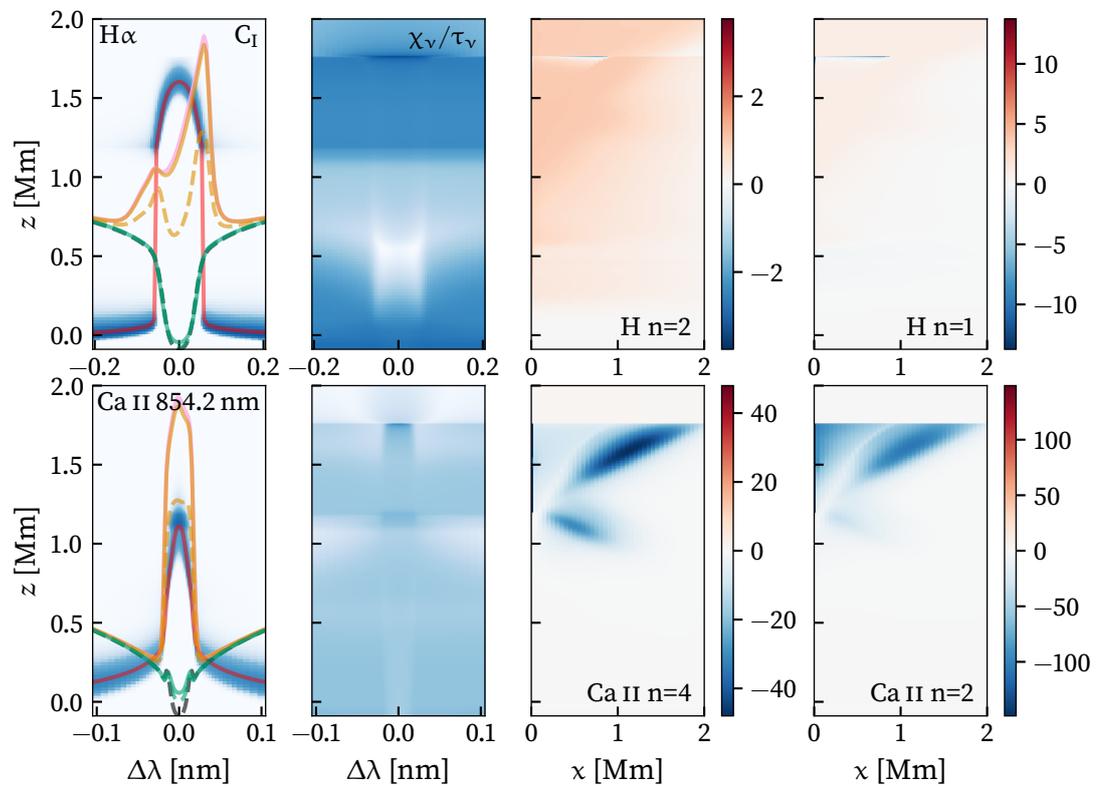
### 6.3 Simulation Results

function, and highlights where the local opacity is large but the optical depth is small. In plane-parallel models with flows this can be used to highlight their effect on the line profile. The line profiles overplotted on the contribution function are the initial statistical equilibrium solution (dashed black), current emission from the flaring boundary (pink), and those observed from 200 (solid line) and 400 km (dashed line) on either side of the flaring kernel (orange for the positive  $x$  region and green for the negative  $x$  region). Whilst it is possible to compute the response function to perturbations in atmospheric parameters, this is not particularly insightful due to the assumption that the temperature and velocity fields remain fixed, and the coupling of the electron density through charge conservation again.

There is an enhancement in the upper and lower levels of the  $H\alpha$  transition in the upper chromosphere appearing deeper in the atmosphere (down to an altitude of 1 Mm) closer to the flaring boundary condition. The contribution function shows that the secondary peaks of the  $H\alpha$  line profile form deeper in the atmosphere than the line core and approximately overlap the bottom of this enhanced region. The depth of this enhancement decreases with distance from the flaring boundary, and corresponds well with the disappearance of the secondary peaks observed over region B. Where the relative change is roughly similar in both the upper and lower levels of the  $H\alpha$  line there will be both an increase in opacity, and an equivalent increase in emissivity. This leads to no net change in the source function, thus this higher altitude region will proportionately contribute more significantly to the outgoing radiation as the increase in opacity blocks light from deeper in the ray's path. This enhancement does not appear to be sufficient to directly create the secondary peaks in the  $H\alpha$  profiles other than in the columns closest to the flaring boundary, as evidenced by the green line profiles, but the increased source function in the higher temperature region of the upper chromosphere is sufficient to significantly enhance the core of the line profile at large distances from the flare ( $\sim 1$  Mm).

The Ca II populations, shown in the lower row of Fig. 6.17, are much more dramatically affected than the hydrogen populations previously discussed. Both the upper and lower levels of the Ca II 854.2 nm spectral line are affected in the same  $z = 1.25 - 1.75$  Mm region at the  $x = 0$  Mm boundary, tapering off as  $x$  increases. The populations are substantially reduced from their initial statistical equilibrium values over this range, with a maximum relative change of approximately  $-20$ . The lower level of this transition is significantly more reduced than the upper level, which will lead to a dramatic reduction in the local opacity whilst the emissivity remains less affected. This will lead to an increase in the local source function as well as a reduction in the opacity affecting the line core radiation formed just below this level. This reduction is likely due to photoionisation effects similar

### 6.3 Simulation Results



**Figure 6.18:** Contribution functions,  $\chi_v/\tau_v$ , and relative changes to atomic populations equivalent to Figure 6.17 for the F10 simulation at  $t = 10$  s.

### 6.3 Simulation Results

to those considered in Sec. 5.4, and will be discussed more in the context of the F10 model.

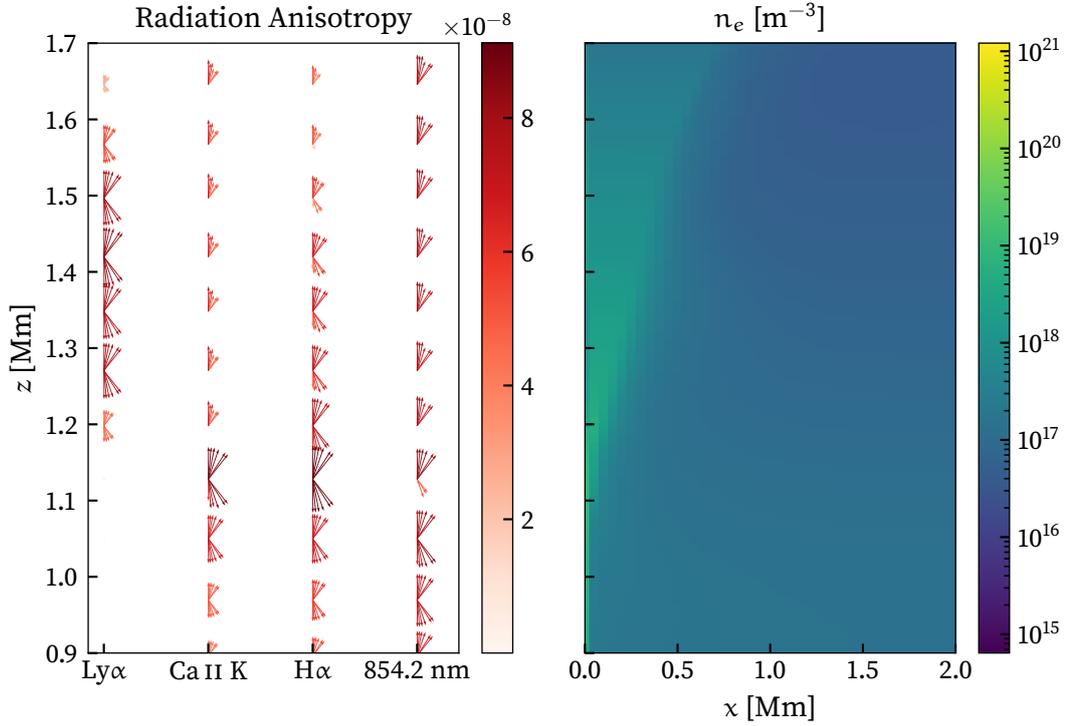
An equivalent figure for the F10 simulation at  $t = 10$  s is shown in Fig. 6.18. Due to the increased magnitude of the population changes in the transition region, it is harder to interpret the population change plots for the F10 model. The upper and lower levels of the  $H\alpha$  transition have been affected similarly to the F9 simulation, but the magnitude of the population increase is larger, however this does not significantly change the interpretation of the population effects on the line profile, which is corroborated by the line profiles shown in the left-most panel. We also note, as evidenced by Fig. 6.16, that the enhancements in the  $H\alpha$  populations are sufficient to produce small secondary peaks in the line profile, and the formation of these is clearer in the contribution function than it was in the F9 case.

Despite the upper level of the Ca II 854.2 nm line appearing darker than the lower level, given the differences in scaling, the lower level populations have again been reduced by a greater amount than the upper level. Thus, the same interpretation as that given to the F9 simulation applies to the Ca II 854.2 nm line over the  $z = 1.25 - 1.75$  Mm region. There is a secondary feature on these Ca II population relative change plots, an additional reduction in the  $z = 1 - 1.25$  Mm region. This feature overlaps the core of the line contribution shown in the left-most panel, and appears to be a relatively similar reduction in both the upper and lower level populations (the reduction in the upper level population becomes larger than that of the lower level population approaching  $x = 1$  Mm).

This will mostly lead to a reduction in local opacity, with a slight reduction in source function. Thus the radiation from deeper in atmosphere, and that formed locally, is less attenuated in both the line-core forming region, and the optically-thick upper-chromospheric layers above the core forming region. This reduction in local opacity appears to be well-validated by the extremely shallow core of the green line-profile in the first panel relative to the statistical equilibrium solution.

An interesting feature present in the Ca II plots is the angled white band starting from  $z = 1.2$  Mm at the flaring boundary and increasing in altitude with distance from the flare. This indicates that there is little change in the calcium populations in this region, despite the much more monotonic variation in  $x$  seen in the F9 case. In Fig. 6.19 the anisotropy of the radiation field leaving the flaring boundary and entering the slab over the range  $z = 0.9 - 1.7$  Mm for the timestep currently considered is shown, along with the electron density in this region of the slab. The anisotropy is shown for the line cores of four different spectral lines ( $Ly\alpha$ , Ca II K,  $H\alpha$ , Ca II 854.2 nm) along the vectors of angular

### 6.3 Simulation Results



**Figure 6.19:** The anisotropy of the radiation field at the flaring boundary at the line cores of four different lines ( $\text{Ly}\alpha$ ,  $\text{Ca II K}$ ,  $\text{H}\alpha$ ,  $\text{Ca II 854.2 nm}$ ) is shown in the left-hand panel for the F10 simulation at  $t = 10$  s. Note that the angles of the rays have been mapped to match the aspect ratio of the right-hand panel. The right-panel shows the electron density in the slab at the same point in time.

quadrature used in the two-dimensional simulation. These have been plotted so that their  $x$  and  $z$  components match the aspect ratio of the electron density plot in the right-hand panel. Their sampling of the unit sphere is far more uniform than it appears here due to this significant stretching effect along  $z$  and each vector's projection in the  $x - z$  plane has been normalised. This mapping allows their directions to be directly considered in terms of the adjacent panel. The length and colour of each vector is then scaled by the associated intensity. In the formation region of each line the line core radiation field is relatively isotropic, but the downgoing rays are greatly reduced above this region (as there is little emission or scattering).

In the right-hand panel there is an approximately triangular region of enhancement in the electron density (which is initially approximately uniform in  $x$  taking the values at the  $x = 2$  Mm boundary), starting at the flaring boundary around  $z = 1.1$  Mm, and increasing

### 6.3 Simulation Results

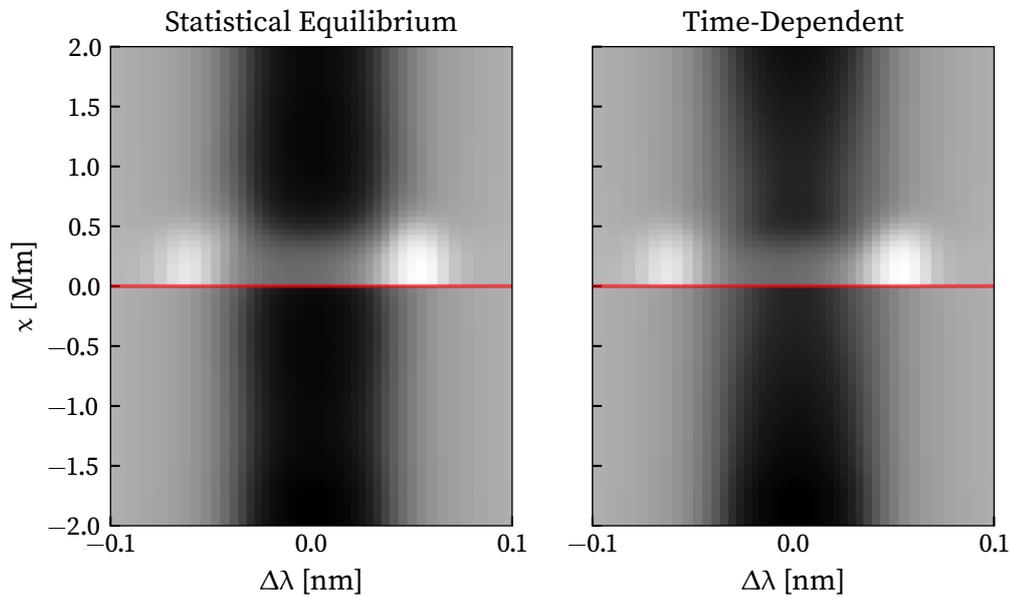
in distance from the flare with altitude up to  $x = 0.75$  Mm at  $z = 1.7$  Mm. The leading edge of this enhancement in electron density agrees well with the white band present in the Ca II population plots and the increase in electron density will locally increase the collisionality of the plasma. This region also appears to overlap well with the incoming Ly $\alpha$  radiation, which will penetrate further at higher altitudes due to the higher temperature and the combination of radiation arriving close to horizontally and from below. This, and other EUV radiation forming at similar temperatures, will likely be the cause of significant photoionisation. The downward sloped population reduction seen in the Ca II population plots corresponds to the line core formation region ( $\sim 1.1$  Mm) which also aligns with the highest altitude at which Ca II 854.2 nm radiation field is relatively isotropic. Above this, the radiation field is primarily upgoing and injects energy into the slab in a non-uniform fashion.

H $\alpha$  appears relatively isotropic up to  $z = 1.45$  Mm, with a slight upward anisotropy likely due to plasma motions of up to  $15 \text{ km s}^{-1}$ , and becoming primarily upgoing above this. From the contribution function in Fig. 6.18 the H $\alpha$  line core in the slab continues to form up to  $z \approx 1.75$  Mm, and the anisotropic injection of energy is likely partially responsible for the diagonal structure seen in the population change plot of H $\alpha$ 's upper level. The Ca II K radiation field behaves similarly to Ca II 854.2 nm, but forms slightly higher, with a strong anisotropic component around  $z = 1.15$  Mm, corresponding roughly with the bottom of the ionisation region depicted in the right-hand panel. This line will be somewhat affected by the region of enhanced electron density, but less than either of the hydrogen lines presented.

There is a significant enhancement of electron density within the triangular region shown in the right-hand panel of Fig. 6.19, which appears to be due primarily to photoionisation (as the temperature within the slab cannot change). This increase in electron density affects line formation within the slab, both in terms of the line profiles, but also collisional rates and LTE populations. It overlaps well with the more complex features shown in the population change plots for the F10 model and shows that the atomic populations responsible for these chromospheric lines can be significantly affected by the incoming radiation.

From the population change figures shown in this section it becomes clear that the far wings and continuum of both of these spectral lines are not affected by the incident radiation due to the lack of changes in the populations in the deep chromosphere and photosphere. It is likely that the plasma here is both too optically thick, and collisionally coupled into LTE to be affected by the incoming radiation. Thus, without a thermodynamic

### 6.3 Simulation Results



**Figure 6.20:** Comparison of statistical equilibrium and time-dependent treatments at  $t = 10$  s for the  $H\alpha$  line in the F9 model.

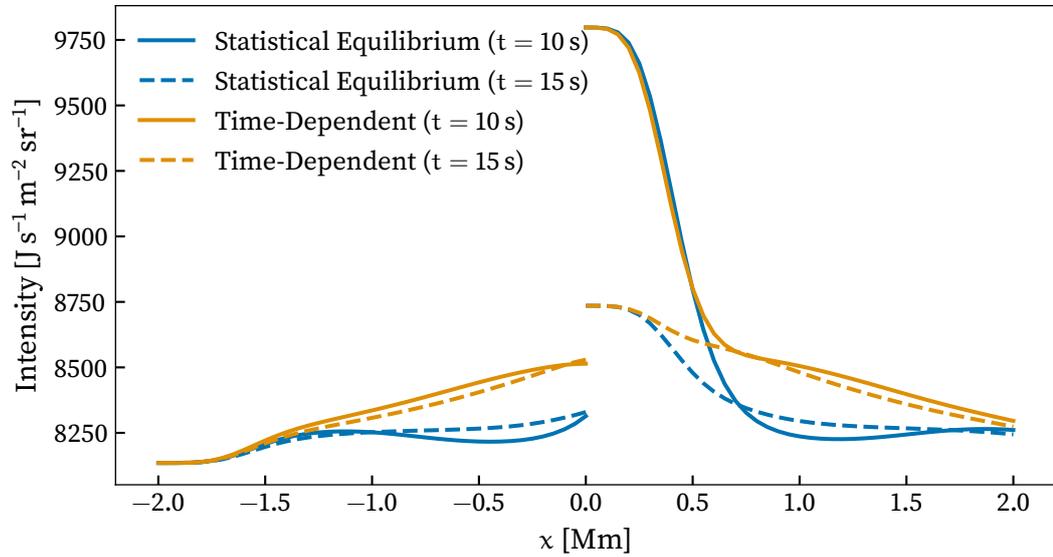
change deep in the slab it is not possible for the continuum levels to change. It is not easy to disentangle how the atomic populations in flaring boundary and the slab both contribute to the outgoing radiation. Nevertheless, it is clear that the flare's effects penetrate far into the slab and have a lasting impact on the populations, on both spatial and temporal scales that should be easily resolvable with upcoming optical instruments.

#### 6.3.4 Importance of Time-Dependence

Due to the need to solve the radiative transfer problem at every timestep in the RADYN simulation, the full time-dependent treatment is extremely computationally demanding. If these effects are not important then the style of simulation presented here can likely be accelerated dramatically by performing a statistical equilibrium solution for the slab illuminated by the flaring boundary at each timestep of interest (which is likely to be many fewer than the total number of timesteps needed for the time-dependent solution).

Statistical equilibrium and time-dependent solutions for the  $H\alpha$  line at  $t = 10$  s in the F9 simulation are shown in Fig. 6.20. Region A appears similar in both cases, but there are significant differences in the line width and core depth outside of this region. It is difficult to directly compare these plots, however the effect is clear when the intensity in this line

### 6.3 Simulation Results

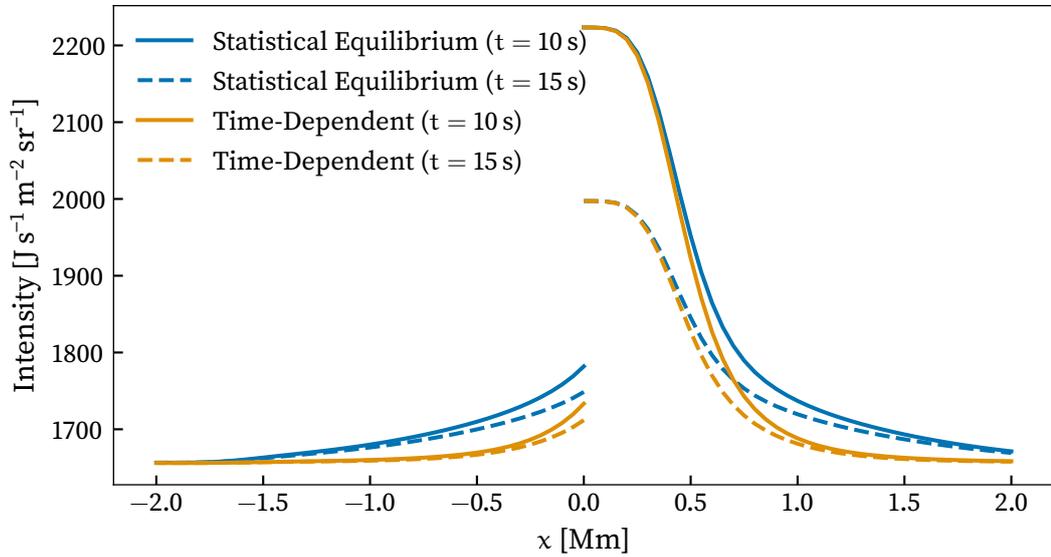


**Figure 6.21:** Integrated intensity of H $\alpha$  lines shown in Fig. 6.20 as a function of  $x$  position.

is integrated and plotted as a function of  $x$  position. This is shown in Fig. 6.21 for both  $t = 10$  s (solid) and  $t = 15$  s (dashed) with the statistical equilibrium treatments shown in blue and the time-dependent treatments shown in orange. The enhancement of H $\alpha$  and the region over which it is enhanced varies substantially between the two treatments, showing that for this line there is a clear need to perform this time-dependent treatment. There is significantly more intensity present in the line in regions B and C, as well as in all of the negative  $x$  region when using the time-dependent treatment. An equivalent comparison is shown for the Ca II 854.2 nm line in Fig. 6.22. The effect is less significant here, but nevertheless easily visually resolvable, and interestingly is the inverse of what is seen for the H $\alpha$  line, as there is less intensity in the time-dependent case than in statistical equilibrium.

The maximum difference between the treatments appears where the primary contribution to the rays is from the two-dimensional slab between a few hundred km and 1 Mm from the flaring boundary. For the positive  $x$  region, the radiation from the boundary mostly dominates region A, but the local conditions within the slab, in particular the electron density, are critical for the line formation along the rays observed from region B. As the flaring boundary is not dominating the outgoing rays from region  $\alpha$ , and these rays look deeper into the slab rather than towards the flaring boundary we immediately see the differences between the treatments close to  $x = 0$  Mm in the negative  $x$  region.

### 6.3 Simulation Results



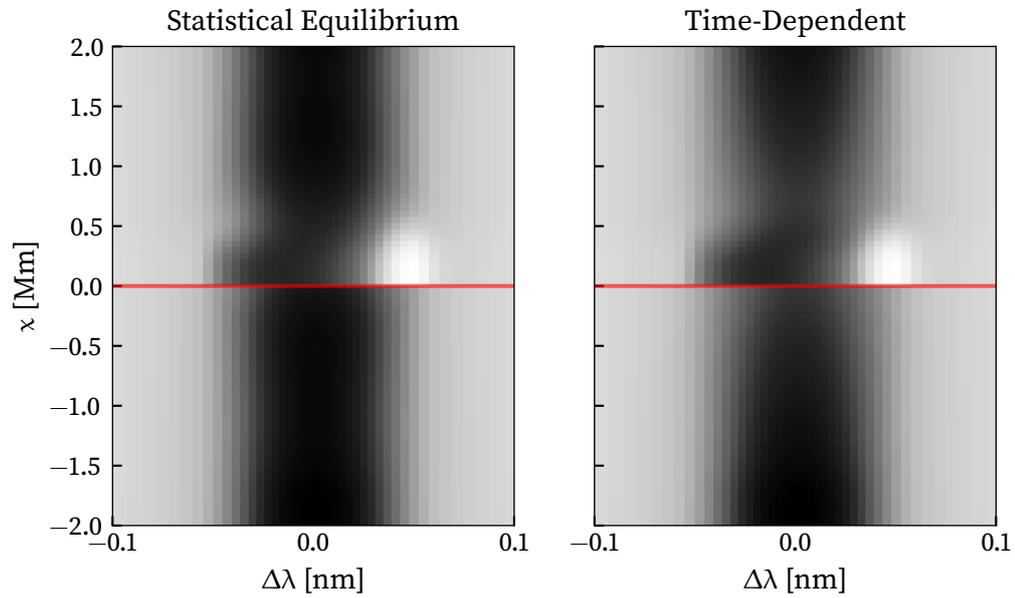
**Figure 6.22:** Integrated intensity of Ca II 854.2 nm lines treated with time dependence and statistical equilibrium as a function of  $x$  position, analogous to Fig. 6.21.

In the wave-heated RADYN simulation of Carlsson & Stein (2002) the settling time for hydrogen to return to equilibrium populations was found to be on the order of  $10^3 - 10^5$  s, much longer than the entire 50 s simulation duration shown here. The effect is likely to be less pronounced here with no thermodynamic variation possible in the slab, and only excitation by incident radiation from the flaring boundary condition, but given the large difference between these treatments at both  $t = 10$  s and  $t = 15$  s it seems prudent to adopt the time-dependent treatment by default when exploring the present phenomenon in flares.

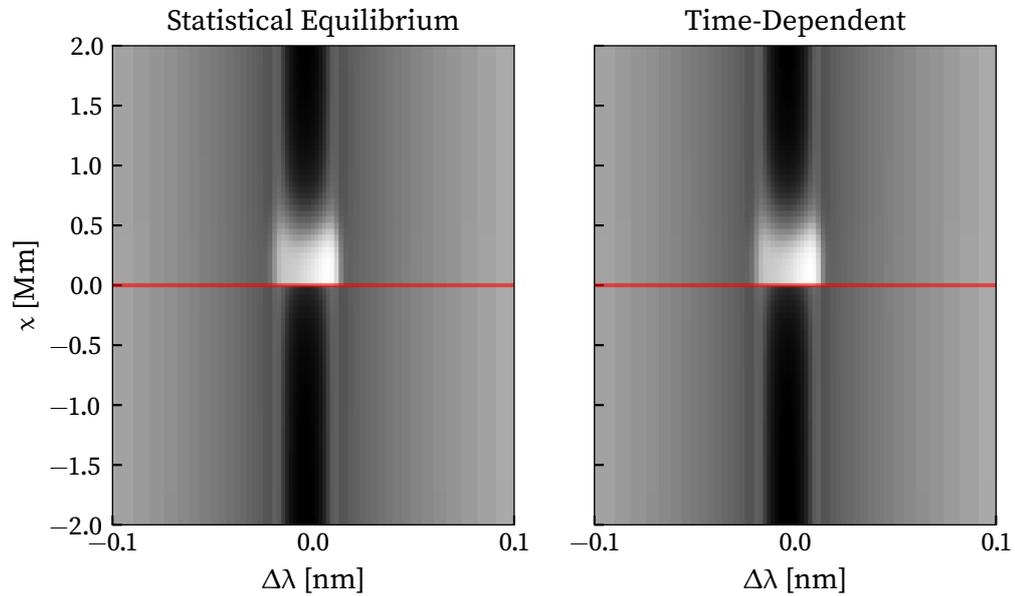
From the discussions in Sec. 5.5 it is reasonable to investigate whether these discrepancies between the time-dependent and statistical equilibrium treatments can be overcome by applying statistical equilibrium with the electron density fixed to the value computed from the time-dependent situation (under the effects of charge conservation).

Figs. 6.23 and 6.24 show the  $H\alpha$  and Ca II 854.2 nm line profiles respectively at  $t = 15$  s using the electron density loaded from the time-dependent model. There remains a significant difference in the  $H\alpha$  line profiles between the two treatments; this is particularly visible in the gradual widening of the line core over regions B and C present in the time-dependent treatment that is barely visible in the statistical equilibrium treatment. We note that the line-core narrowing around  $x = 0.5$  Mm in this statistical equilibrium case is not present in

### 6.3 Simulation Results



**Figure 6.23:** Comparison of statistical equilibrium and time-dependent treatments with electron density taken from time-dependent model at  $t = 15$  s for the H $\alpha$  line in the two-dimensional slab case.



**Figure 6.24:** Comparison of statistical equilibrium and time-dependent treatments with electron density taken from time-dependent model at  $t = 15$  s for the Ca II 854.2 nm line in the two-dimensional slab case.

### 6.3 Simulation Results

Fig. 6.20 and must be due to the use of the electron density from the time-dependent model. The  $H\alpha$  line profile is therefore closer to that expected from the time-dependent case, but significant differences remain, and it appears that a full time-dependent treatment is necessary to capture the evolution of the hydrogen populations.

Ca II 854.2 nm, on the other hand, agrees well with the expected solution when computed in statistical equilibrium with the provided electron density. This is not surprising following the analysis undertaken in Sec. 5.5, but we note that this is only one timestep of the model.

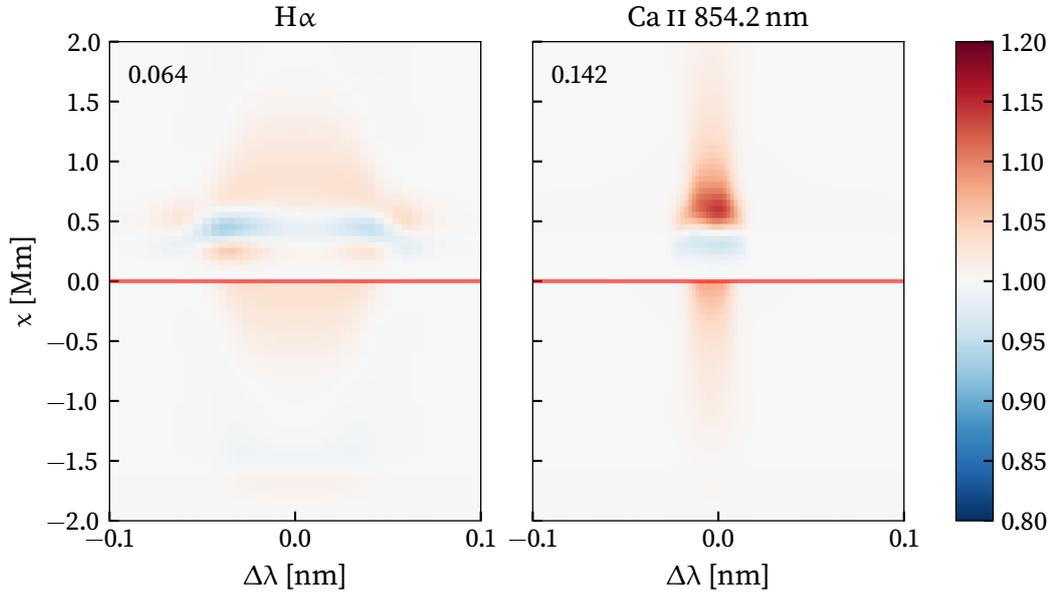
Attempts were made to compare these treatments at other timesteps but the process of converging a statistical equilibrium solution with the fixed electron density provided from the time-dependent model proved to be unreliable, and the models at  $t = 5, 10,$  and  $11$  s failed to converge. This approach therefore appears unsuited to the models presented here, but from the comparison of the Ca II 854.2 nm line profiles and the analysis undertaken in the previous chapter it is plausible that a situation with a detailed static treatment of hydrogen, where the populations and electron density can be known approximately, could produce relatively accurate Ca II 854.2 nm line profiles, at lower computational cost than the full time-dependent method employed in this chapter.

#### 6.3.5 Horizontal Resolution

Whilst it would be nice to evaluate the radiation with arbitrary horizontal precision, the limitations of computer performance do not allow this. Clearly, we want to investigate the problem at a resolution similar or better than that provided by our observations. If the problem produces smoothly varying solutions then it is not necessarily required to run the simulation at the same resolution as the observations; however we anticipate fine features that only span one or two pixels, and thus use a 50 km element spacing as our baseline. This is similar to the 42 km resolution provided by CRISP observation against which our results will be compared.

To analyse whether this resolution appears sufficient we shall compare the statistical equilibrium solution to our problem (as this only requires the calculation of a single timestep, rather than the entire model) at a horizontal grid spacing of both 50 km and 25 km at several timesteps in the F9 simulation and compare the outgoing radiation. It is reasonable to use a statistical equilibrium treatment for this as we are primarily testing the performance of the formal solver on this atmospheric configuration, and it

### 6.3 Simulation Results



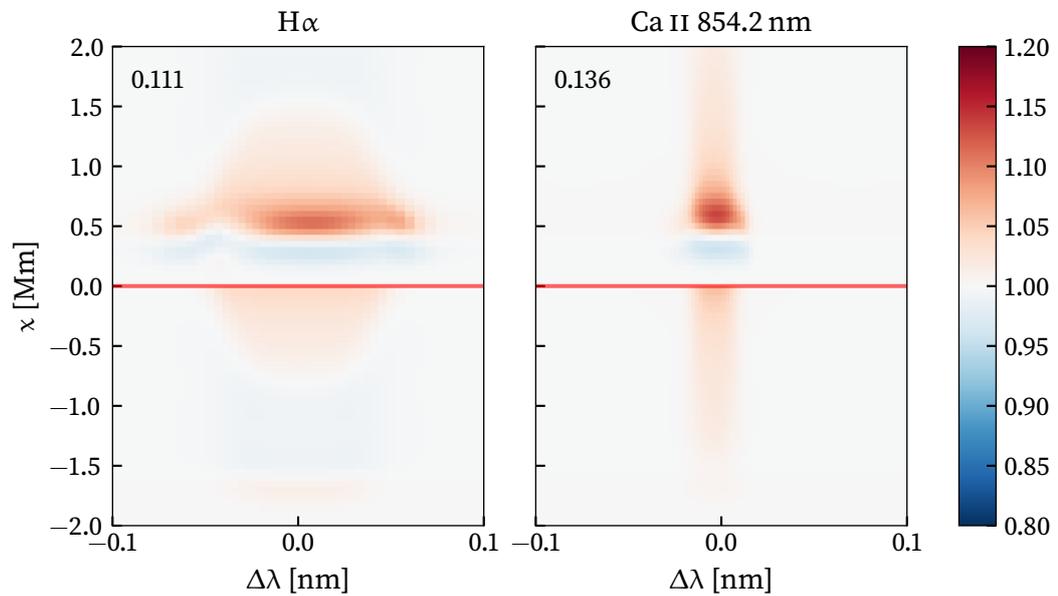
**Figure 6.25:** Ratio of outgoing statistical equilibrium radiation using 41 and 81 horizontal points in H $\alpha$  (left) and Ca II 854.2 nm (right) at  $t = 5$  s. The number in the upper-left corner is the maximum change in intensity in each panel (i.e. the absolute value of the ratio presented in the figure subtracted from unity.)

should converge to similar solutions in both cases if sufficient resolution is used. The incident radiation at the  $x = 0$  Mm boundary is the same in both instances and the atmospheric parameters are horizontally uniform, although the electron density in the slab is allowed to vary through charge conservation. In Figs. 6.25–6.27 we show the ratio of the outgoing radiation from the 41 point (50 km) and 81 point (25 km) models at  $t = \{5, 10, 15\}$  s respectively. The maximum fractional error is shown in the upper-left corner of each panel.

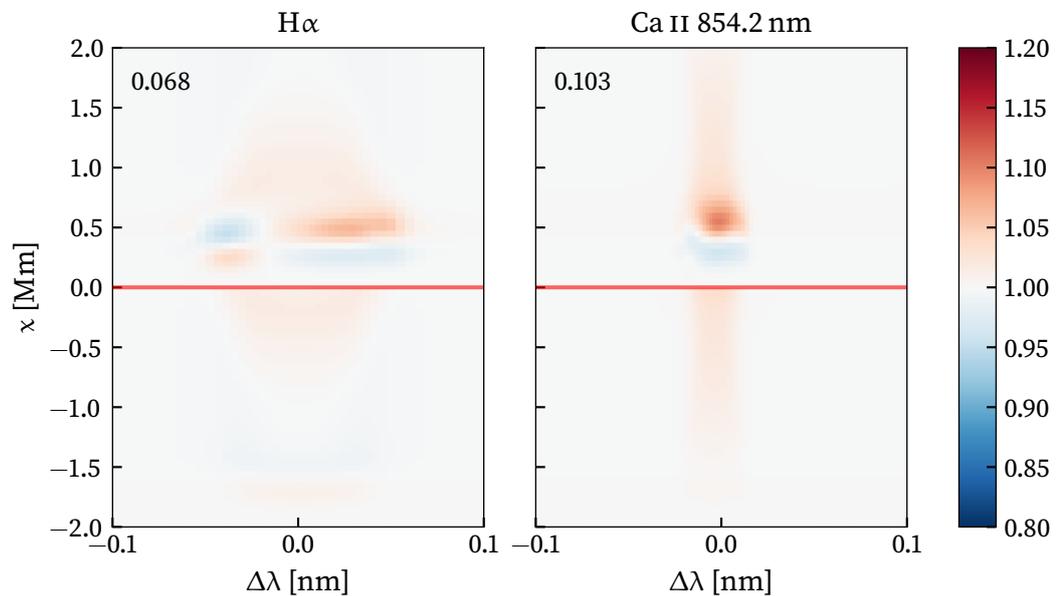
The location of the peak differences between the treatments in the positive and negative  $x$  regions appear to follow a pattern similar to those considered in Sec. 6.3.4, with the largest differences in positive  $x$  appearing in region B while those in negative  $x$  appear in  $\alpha$ . The maximum difference is found to be an enhancement of around 14% in both line cores at  $x \approx 0.6$  Mm for the ratio of lower resolution simulation to the higher resolution. This is not an insignificant variation, but it does not appear to vary dramatically during the simulation, and the variation is smooth, with no significant effects on the line shape.

The computational cost of using the higher resolution model is unfortunately not a simple linear ratio of the number of points (i.e. the 81 point simulation does not take simply  $\sim 2\times$  longer than the 41 point simulation). Indeed, the cost of each iteration scales close to

### 6.3 Simulation Results



**Figure 6.26:** Ratio of outgoing statistical equilibrium radiation using 41 and 81 horizontal points in H $\alpha$  (left) and Ca II 854.2 nm (right) at  $t = 10$  s. The number in the upper-left corner is the maximum change in intensity in each panel (i.e. the absolute value of the ratio presented in the figure subtracted from unity.)

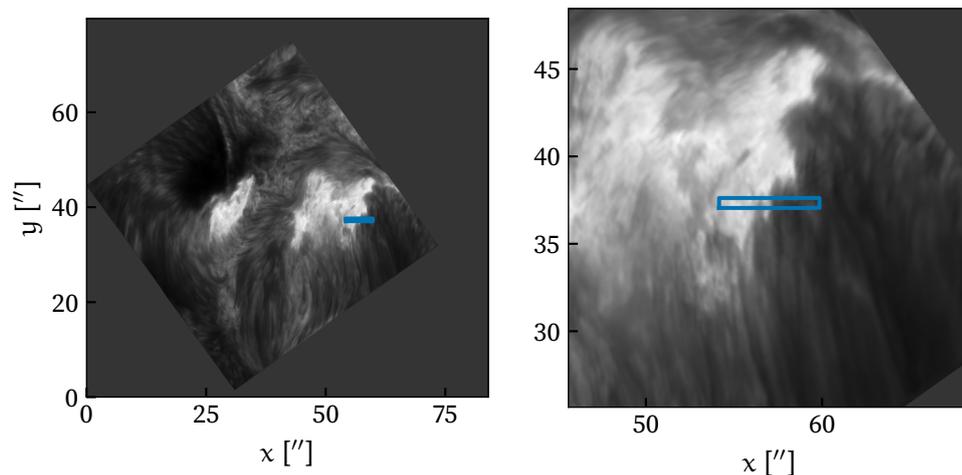


**Figure 6.27:** Ratio of outgoing statistical equilibrium radiation using 41 and 81 horizontal points in H $\alpha$  (left) and Ca II 854.2 nm (right) at  $t = 15$  s. The number in the upper-left corner is the maximum change in intensity in each panel (i.e. the absolute value of the ratio presented in the figure subtracted from unity.)

## 6.4 Observational Comparison

linearly with the number of points in the atmosphere (for a given angular and wavelength quadrature), but the number of iterations needed also increases with finer grid sizes, as the information carried by the approximate  $\Lambda$  operator takes longer to cross the grid. For this reason the 81 point simulation takes  $\sim 2.4\times$  longer than the 41 point simulation. We therefore opted to use the 41 point spatial grid in all models presented in this chapter, primarily for computational reasons, but are confident that the solution obtained is reasonable.

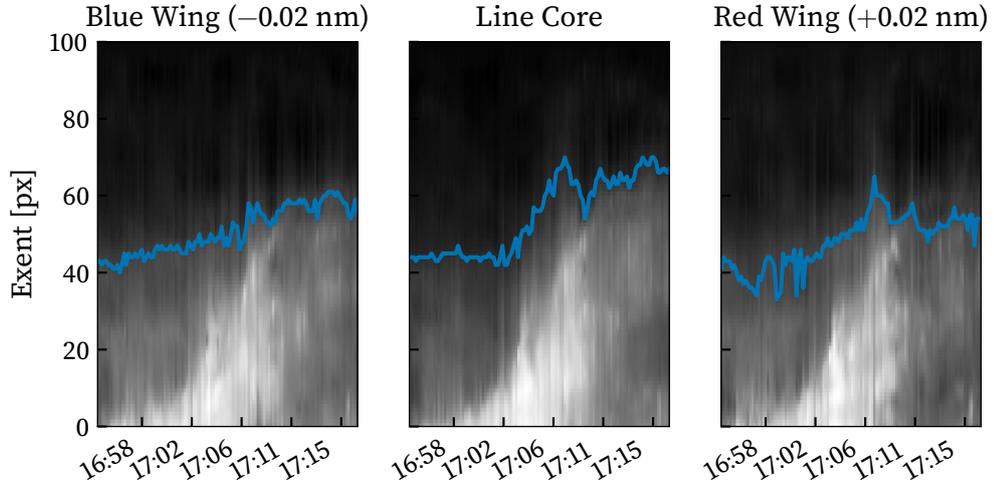
### 6.4 Observational Comparison



**Figure 6.28:** Ca II line core observation of M1.1 flare at time 2014-09-06T17:09:26

Whilst any real world flare situation is far more complex than the simplified two-dimensional model presented in the previous section, we will compare these results against observations of the two-ribbon M1.1 flare SOL 20140906T17:09 from the CRISP instrument, observed in  $H\alpha$  and Ca II 854.2 nm. This data was prepared following the standard procedure described in Sec. 3.1.1. A context image of the region in the core of the Ca II 854.2 nm line at the flare peak is shown in Fig. 6.28. We have selected a region on edge of the western ribbon for this comparison, which is highlighted in blue in both the left- and right-hand panels of this figure. The chosen region is  $100 \times 10$  px, with each pixel of the CRISP instrument being  $0.057''$  ( $\sim 42$  km). This region was chosen as it represents the

## 6.4 Observational Comparison



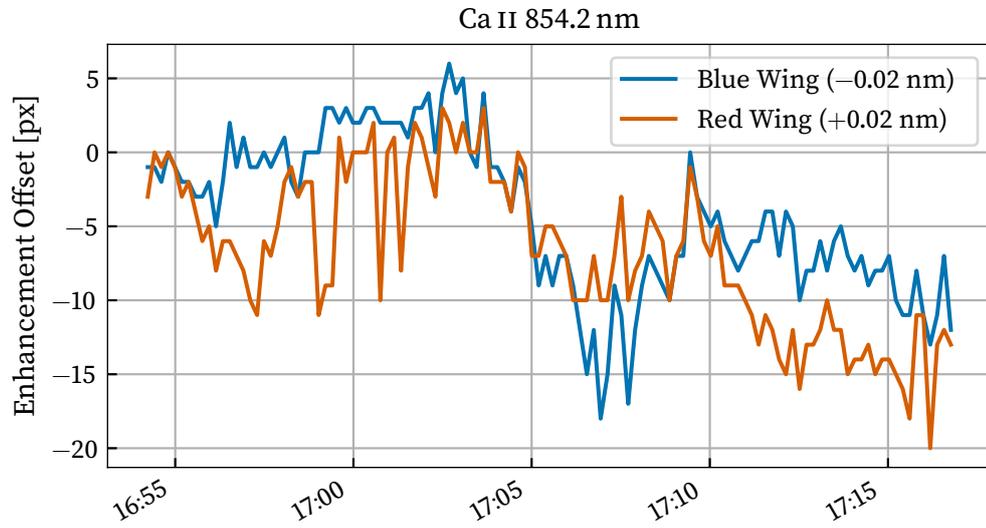
**Figure 6.29:** Extent of enhancement in Ca II 854.2 nm line for the region labelled in Fig. 6.28.

intersection between the flare ribbon and some of the quieter material present in the data (the east ribbon overlaps a sunspot), lies approximately perpendicular to the structure visible here, and is close to the region analysed by Kuridze et al. (2015). We note that for both spectral lines, the angular size of 2 px is less than the diffraction limit of the SST at these wavelengths (Sec. 3.1.1), and this may slightly affect the extents of the enhancements measured in this data.

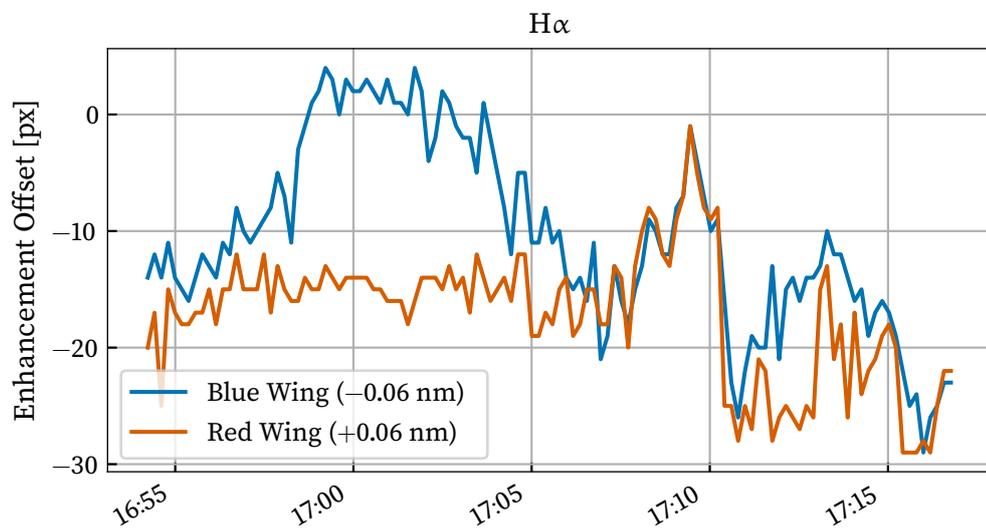
To investigate the effects presented in the simulation we shall analyse the variation of the radiation in the selected region as a function of  $x$  location and time. The 10 px in  $y$  are averaged to increase the signal-to-noise ratio. These timeslices are then plotted for three different wavelengths, corresponding to the two wings and the line core, and shown in Fig. 6.29. We take the right-hand edge of this region as representing an approximation of the background intensity and compute the distance from the left-hand boundary at which a 40% enhancement over this background level is seen. This extent is shown as the blue lines in Fig. 6.29. The bright expansion of the flare ribbon can be seen on these plots as the triangular region that collapses rapidly around 17:11. This figure is similar in concept to the imaging plots (Figs. 6.8–6.11) produced from our models, where the “ribbon” was held at the  $x = 0$  boundary.

We can compare the differences between the extents for these different spectral positions on both lines to investigate the differences in emission throughout the slab. This is shown for Ca II 854.2 nm and H $\alpha$  in Figs. 6.30 and 6.31 respectively. The comparison here is against

#### 6.4 Observational Comparison



**Figure 6.30:** Offset between the enhancement of the blue and red wings of the Ca II 854.2 nm line and the line core for the region labelled in Fig. 6.28.



**Figure 6.31:** Offset between the enhancement of the blue and red wings of the H $\alpha$  line and the line core for the region labelled in Fig. 6.28.

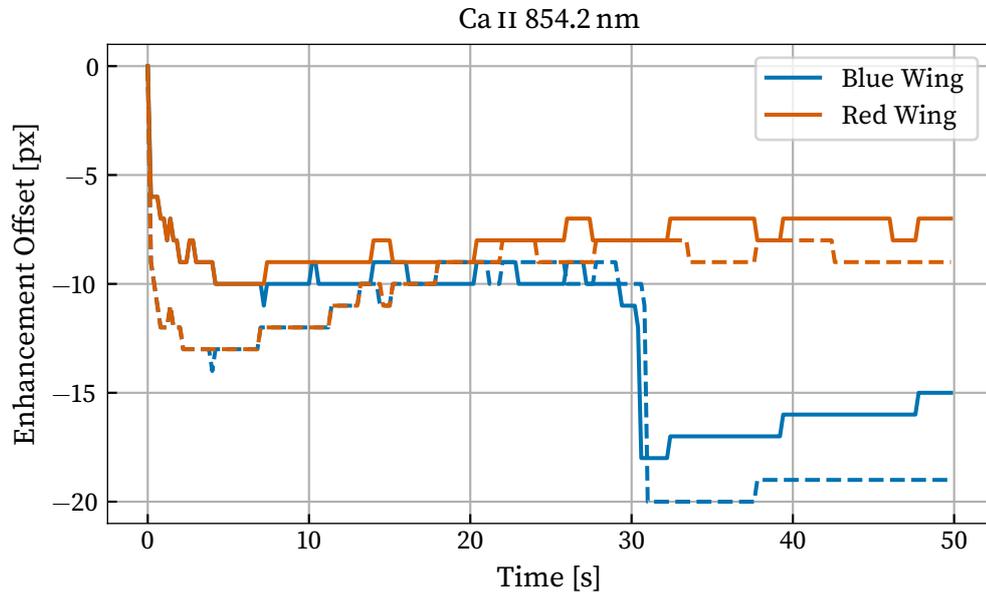
## 6.4 Observational Comparison

the line core, so an offset of 0 px implies that the enhancement is co-spatial in the line core and at this wavelength. For Ca II 854.2 nm, prior to 17:05, there is little difference between both line wings and the line core, although the blue wing enhancement on average extends slightly further than the red wing, and just prior to the flare peak, up to 5 px further than the line core. As the flare peaks, the enhancement of the line core extends consistently 5 – 10 px further from the ribbon than either of the wings, and this trend continues for the remainder of the dataset. For H $\alpha$ , before  $\sim$ 17:07, the red wing enhancement remains  $\sim$ 15 px behind the line core, whilst the blue wing starts in this region, then extends past the line core at  $\sim$ 17:00 before falling back to a similar offset to the red wing at  $\sim$ 17:07. From this point onwards, both wings remain at a similar offset, typically remaining 10 – 30 px less extended than the line core.

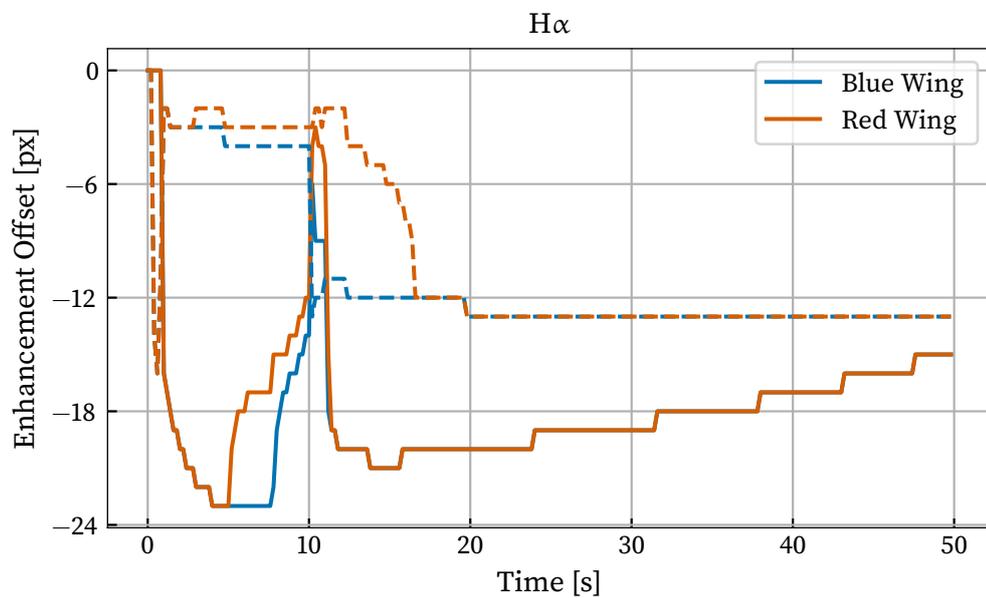
We can compute this same metric for the simulations of slabs adjacent to F9 and F10 flares presented in this chapter, and these are shown in Figs. 6.32 and 6.33 respectively for the Ca II 854.2 nm and H $\alpha$  results. The red and blue wings are again shown with the same colours, the solid lines refer to the F9 simulation and the dashed lines the F10 simulation. We note that there is a substantial variation in viewing angle between these simulations and the set of observations investigated here. The simulations use a viewing angle  $\mu_z \approx 0.951$  (an inclination of  $2.24^\circ$  when projected into the  $x - z$  plane), whilst this observation was taken at  $\mu \approx 0.565$ . If the same viewing angle is used with the simulation then all of the “pixels” at the top of the atmosphere “look” directly into the flaring boundary condition, and the effect we wish to investigate cannot be resolved without performing simulations with substantially larger extents in  $x$ . Due to the prohibitive computational cost of this (a similar horizontal spatial resolution would still be required), we instead choose to compare the simulations using this smaller viewing angle as a “first-order approximation” of the problem. In both cases, the size of a pixel is similar (42 km for CRISP and 50 km in the simulation); whilst the pixels in the data are nominally smaller this does not include the effects of viewing the inclined solar surface so we compare these extents in pixels directly.

The Ca II 854.2 nm line behaves similarly in both the F9 and F10 simulations, with the line core immediately extending significantly further than the wings (by 9 – 13 px) as the heating begins. Around  $t = 30$  s the extent of the blue wing enhancement decreases significantly, whilst the red wing enhancement remains approximately constant, slightly approaching the line core. This is not in agreement with the observations where the red wing offset decreases relative to the blue wing offset later in the cooling phase. The consistency of the red enhancement is likely due to the strong red asymmetry present in region A of the simulation at later times, whilst the core remains notably enhanced

## 6.4 Observational Comparison



**Figure 6.32:** Offset between the enhancement of the blue and red wings of the Ca II 854.2 nm line and the line core. The dashed lines show the F10 simulation, and the solid lines the F9.



**Figure 6.33:** Offset between the enhancement of the blue and red wings of the H $\alpha$  line and the line core. The dashed lines show the F10 simulation, and the solid lines the F9.

#### 6.4 Observational Comparison

further into the slab. Nevertheless, the extent of the line core enhancement relative to the wings remains comparable to the observation, in the 5 – 20 px range.

There are significant differences between the F9 and F10 simulations for the H $\alpha$  line. In the less energetic simulation, the line core offset rapidly jumps to  $\sim 22$  px further than both the red and the blue wings. As the heating ends there is a short-term variation (lasting  $\sim 1$  s) in the red and blue wing extents, as both approach the line core extent. For the remainder of the cooling period, the line core remains 15 – 20 px more extended than both of the line wings. This is approximately comparable to the observed offset from 17:10 onwards (10 – 30 px), and the short-term spike after the heating ends could be compared to the spike in the observation just prior to 17:10, although the timescales are very different. For the F10 simulation, the difference in extent between the line core and the wings remains small during heating, but after the heating ends increases relatively rapidly to  $\sim 13$  px. The red wing takes approximately 8 s longer to drop back to this level.

Qualitatively, there are a number of features that agree between the simulations and this observation: from the flare peak onwards both line cores remain significantly more extended than their respective wings, and both simulations support this, with comparable changes in enhancement offset. After the flare peak, the line-core remains visibly extended into the quieter region adjacent to the western ribbon, up to 20 px further than an equivalent enhancement in the wings. In both the simulation and the observation this is in part due to the variation of the line profile shape, changing from either a strong emission or centrally reversed enhanced line on the ribbon to a much deeper absorption line off-ribbon. Given the additional complexity of the observed configuration, the difference in viewing angle, timescale, and the likelihood of thermodynamic changes occurring in proximity to the flare (including Doppler shifts from plasma flows), it is unwise to attempt to interpret the comparison of these offsets in much greater depth than this qualitative agreement. The basic agreement between the simulation and observation suggests that this simulation is an accurate “first-order” approximation of this problem, and is capturing some of the observed effects through the treatment of radiation alone, although further simulation of more complex configurations and in-depth comparisons are needed.

Future comparisons between this style of modelling and observations can follow the methodology presented here, although it would be advantageous to more closely match the viewing angle of the observation and simulation. Whilst it is important to consider inclined observations, when using the approach presented here, slightly inclined simulations are significantly less computationally intensive than ones with extreme inclination, due to the need for the rays exiting the top of the atmosphere to not simply sample the flaring

boundary. This can likely be addressed by modifying the functionality of the formal solver for the final synthesis, so as to pick rays that will intersect with the chromospheric regions of interest.

## 6.5 Discussions

The basic two-dimensional simulations shown in Sec. 6.3 demonstrate that even for a two-dimensional slab of plasma with a fixed quiet Sun temperature structure, the time-dependent illumination by an adjacent flare will create significant variations in the level populations deep in the slab, causing notable enhancements in the observed  $H\alpha$  and Ca II 854.2 nm optical spectral lines 1 Mm and further from the flaring boundary. Traditional inversion techniques performed on this region are likely to lead to incorrect conclusions as to the thermodynamic structure of the slab, as the outgoing radiation is not necessarily directly linked to the slab's local thermodynamic properties, but is instead significantly influenced by the incident radiation. If the thermodynamic properties were allowed to vary through absorption of radiation and cross-field heat conduction (it is difficult to foresee other terms being possible without resorting to 2D RMHD simulation), then the local thermodynamic properties would be more closely linked to the observed radiation and changes in continuum flux would become possible. We have shown that the effects of radiation alone in the chromosphere are sufficient to produce significant changes in the observed line profiles and it does not seem likely that thermodynamic effects would dominate given the magnitude of the radiative effects, so care should still be taken when applying traditional inversion techniques.

It is necessary to perform a time-dependent treatment of the hydrogen populations in this problem, taking into account the effects of non-equilibrium ionisation, and this species' slow return to statistical equilibrium, despite the lack of time-varying thermodynamic modifications of the quiet Sun slab. With the electron density provided by non-equilibrium hydrogen ionisation, a statistical equilibrium treatment of calcium was able to accurately reproduce the time-dependent Ca II 854.2 nm line profile (similarly to our results from Chap. 5). We note that this result is valid only for one timestep of this model, as it was the only one that converged, suggesting the importance of the full time-dependent treatment.

The enhancement effects are not uniform as a function of wavelength, with the line core enhancement typically extending up to several hundred km further than the enhance-

## 6.5 Discussions

ments in the line wings. This may have implications for the calculation of filling factors, due to the wavelength dependence of this “smearing”, but more detailed simulation and analysis are needed to attempt to prescribe any functional form for this.

Nevertheless, these enhancements are on scales that are observable with current generation instruments such as SST/CRISP, and are likely to become more important with the improved resolution of upcoming telescopes such as DKIST. We have presented an initial analysis of these effects in SST/CRISP data and found enhancements of a similar order of magnitude sharing a number of observational characteristics, but significant differences remain, likely due to the limitations of our simple model. It is likely that the temperatures in such a slab will vary due to the absorption of radiation, thermal conduction and other modes of heating occurring adjacent to the flare. Thus, the results presented here likely underestimate the magnitude of these line enhancement effects which could be more accurately captured by using the detailed two-dimensional radiative transfer available in *Lightweaver* in conjunction with an RMHD model responding to the incoming energy.

# 7

## RADYNVERSION

*The contents of this chapter are based on my contributions to the research presented in Osborne et al. (2019).*

As discussed in Sec. 3.2, the assumptions that render NLTE response function based approach to inversions tractable, such as hydrostatic equilibrium, cannot reasonably be applied in flares where flows often approach the sound speed. We therefore need an inversion technique that can operate outside of these constraints. In the notation of Sec. 3.2, our standard radiative transfer forward process can be framed as a function  $y = f(x)$  with atmospheric inputs  $x$  and line profiles  $y$ . Clearly this function is not bijective, but if we also capture the information lost in the forward process, we can instead define a bijective function  $x = g(y, z)$  such that  $g^{-1}$  represents the forward process, and  $g$  the inverse process. Our theory of radiative transfer does not give any immediate insight into the formulation of  $g$  with so few constraints, so we instead turn to the field of machine learning from RHD simulations with the intention of learning the form of the information lost in the forward process.

The RADYNVERSION model was the first machine learning model for solar inversions to address the problem of degeneracies in the inversion, but not the first to approach inversions of solar observations. Simple fully connected networks have long been used to obtain plausible atmospheric parameters based on Milne-Eddington inversions (e.g. Carroll & Staude 2001; Socas-Navarro 2005). Asensio Ramos & Díaz Baso (2019) and Milić & Gafeira (2020) both presented models focused on the photosphere, that learn to mimic the results of traditional inversions with convolutional neural networks, the former using spatial correlations and allowing inference of the depth of the Wilson depression, the latter using a much simpler model and taking a pixel-by-pixel approach. These models are not theoretically capable of handling degeneracies, but in practice produce plausible results, much faster than is possible with any traditional gradient-based inversion code.

## 7.1 The RADYNVERSION Model

Gafeira et al. (2021) caution against the direct use of these CNN models directly, and instead applied the model of Milić & Gafeira (2020) to initialise traditional inversions, allowing them to converge substantially faster, and in some cases, produce more accurate fits. More recently, Díaz Baso et al. (2021) have presented an elegant Bayesian method using normalising flows, a family of methods related to the one we present below. They present the results of this method and show that it is capable of providing plausible inversions and posterior distributions using observations of Ca II 854.2 nm and Fe I 630.15 nm. The field of machine learning based, and assisted, inversions is undergoing rapid development and appears to have a very exciting future.

The RADYNVERSION model and associated training methods are available under the permissive MIT license on GitHub<sup>1</sup> with archival on Zenodo (Osborne & Armstrong 2019).

### 7.1 The RADYNVERSION Model

Our approach to inversions of solar flares is based upon the application of machine learning to the problem as it is framed in Fig. 3.3. We wish to learn the form of the latent space, and by sampling this space sufficiently for a given observation we can infer the likelihoods of parameters at each location in the flaring atmosphere. Nevertheless, we have no data on which to train which directly characterises the latent space, and thus we turn to the technique of invertible neural networks (INNs) which naturally learn bijective functions to learn the bijective mapping *and* the form of the latent space simultaneously.

The INN is a form of DNN in which invertible blocks are used. A traditional fully connected or convolutional layer used in an ANN is not generally invertible. Indeed, whilst a fully connected layer with equal number of inputs and outputs and an invertible activation function *can* be inverted, it is extremely computationally expensive to do so due to the cost of inverting the potentially large square matrix of weights. The INN is instead built on trivially invertible blocks which can be trained to naturally learn our bijective function. These blocks are built using DNNs, and are thus universal function approximators themselves, meaning they can learn to map different density (in the sense of probability *density*) distributions to each other. This is the cornerstone of how the form of the latent space is learned: the INN is trained to map the samples of the latent distribution to a known distribution. Provided there is sufficient dimensionality to represent the original distribution, it can be remapped to another without any loss of information. We can

---

<sup>1</sup><https://github.com/Goobley/Radynversion>

## 7.1 The RADYNVERSION Model

therefore choose to represent the latent space as any smooth continuous distribution desired and the INN will internally learn the mapping to the form of the “true” latent space contained within the training data. For simplicity we choose to represent the latent space as the unit multivariate Gaussian distribution with mean 0 and variance 1, which we denote  $\mathcal{N}(0, \mathcal{I}_n)$  for an  $n$  dimensional case. Due to the speed of our neural network based solution relative to the cost of the formal solution and iteration that is required in conventional NLTE inversions, we can simply take a large number of samplings of our chosen latent space, and thus generate a probability density function for the atmospheric parameters at each location in the atmosphere, conditioned by our training set. The approach to inferring atmospheric parameters using the INN therefore differs to the approach of traditional methods that take guided explorations of the latent space by taking much broader samplings of this space and constructing a posterior distribution from these, but similarities can also be drawn between this rapid direct sampling and the “exploration” approach of traditional inversion codes. Pre-trained models such as INNs can have significant performance advantages when applied to many similar inputs, as regression-based approaches will tend to treat these independently and repeat similar trajectories, whereas an INN’s training is shared between each application.

The invertible blocks on which our INN is built are known as *affine coupling layers* (Dinh et al. 2014, 2016) of the form first presented by Ardizzone et al. (2018). The input vector  $\vec{x}$  is first split into two halves  $[x_1, x_2]$  which undergo the following affine transformations

$$y_1 = x_1 \otimes \exp(s_2(x_2)) + t_2(x_2), \quad (7.1)$$

$$y_2 = x_2 \otimes \exp(s_1(y_1)) + t_1(y_1), \quad (7.2)$$

where  $\otimes$  represents the elementwise product of tensors and  $s_i$  and  $t_i$  ( $i \in \{1, 2\}$ ) are arbitrarily complex differentiable functions (that need not be invertible). It is worth noting the order of operations here, as  $y_1$  must be computed before  $y_2$ . The output  $\vec{y}$  is then constructed from the concatenation of  $y_1$  and  $y_2$ . The inverse of these affine transforms is given by

$$x_2 = (y_2 - t_1(y_1)) \oslash \exp(s_1(y_1)), \quad (7.3)$$

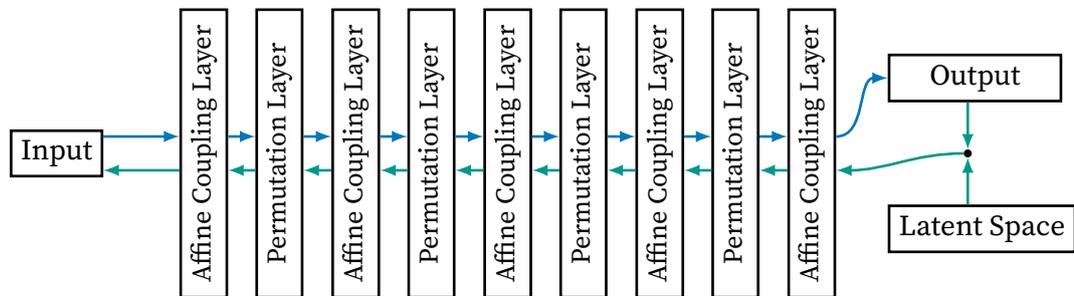
$$x_1 = (y_1 - t_2(x_2)) \oslash \exp(s_2(x_2)), \quad (7.4)$$

where  $\oslash$  represents the elementwise division of tensors. We have stated that the  $s_i$  and  $t_i$  functions can be of arbitrary complexity, but clearly they need to be tailored to the particular task and for this reason we apply DNNs in this role. The networks used for  $s_i$

## 7.1 The RADYNVERSION Model

and  $t_i$  are identical, save for the application of a  $\tan^{-1}$  transformation on the output of the  $s_i$  block, which prevents extreme values from being produced whereby the exponential term would dominate, or have no effect on the output of the block.

In the spirit of DNNs we then stack multiple affine coupling layers to allow for increased representational capability within the INN. As the input is split in half upon entering each affine layer, we can see that the data in each half is only combined at the elementwise multiplication step in each layer on its journey through the network. To alleviate this, and further increase the generalisation capabilities of the network, we interleave a permutation layer between each affine coupling layer. This layer shuffles the data in a random, but fixed order which is different for each permutation layer, before it is split, whilst allowing for trivial reversibility.



**Figure 7.1:** Structure of the final RADYNVERSION network. The forward process is shown with the blue arrows, and the inverse process with teal arrows.

For the RADYNVERSION<sup>2</sup> model we use five affine coupling layers, with four interleaved permutation layers. This is shown schematically in Fig. 7.1. Given the presence of four DNNs per affine coupling layer, our final model is composed of twenty DNNs. Each of these networks is an individual four-layer fully connected network utilising leaky ReLU activation functions after each of the first three layers, and a ReLU following the final layer. This architecture was developed empirically, guided both by the examples presented in (Ardizzone et al. 2018) and machine-learning best practices. A commonly used approach that we applied extensively is to take a small subset of the training data and attempt to find a model which can successfully overfit this mapping; a model that can memorise the mapping for a subset of inputs can demonstrably learn the problem (although it may not have enough dimensionality to generalise, depending on the range and complexity of the input data). Models with both three and five affine coupling layers were trialled, with the latter performing significantly better. Whilst it is possible to continue to increase

<sup>2</sup>RADYNVERSION is a portmanteau of RADYN and “inversion”.

## 7.1 The RADYNVERSION Model

the depth of this network, this is not without significant computational cost, and did not seem necessary given the quality of the results.

RADYNVERSION is trained using RHD simulations generated with RADYN, which contain the structure of a model flaring atmosphere and the synthesised emergent radiation (here the H $\alpha$  and Ca II 854.2 nm spectral lines). The input to the RADYNVERSION network then consists of a reduced set of atmospheric parameters at a given point in time; in our case the temperature  $T$ , electron density  $n_e$ , and velocity  $v$ . These are provided to the network on a fixed height stratification, with 50 points covering the entire atmosphere. The large dynamic range of these parameters (both across a single atmosphere, and between different timesteps at a fixed location in the atmosphere) can have a negative impact on the training and accuracy of the ANN, and we therefore map  $T \mapsto \log_{10} T$ ,  $n_e \mapsto \log_{10} n_e$ , and  $v \mapsto \text{sign}(v) \log_{10} (|v_{\text{km s}^{-1}}| + 1)$ . The mapping for  $v$  serves to scale it based on its base-10 logarithmic value, whilst preserving its sign information. These logarithmic mappings preserve detail better over each decade than a simple linear rescaling.

The output consists of line profiles, in this case H $\alpha$  and Ca II 854.2 nm. Both of these are interpolated onto fixed wavelength grids, with 30 points each (a half width of 0.14 nm for H $\alpha$  and 0.1 nm for Ca II 854.2 nm). The intensity values of the spectral lines are scaled to cover the range [0, 1] whilst preserving the relative intensity of the two (which conveys important information regarding continuum emission).

Our model then has an input dimensionality of 150 and an output dimensionality of 60. We choose, by experimentation, to set the size of the latent space to the same as the input, however we cannot prove the optimality of such a choice as it depends on the (unknown) intrinsic dimensionality of the problem. Thus the input and total output (i.e. output and latent space) must be of length at least 210. To improve the generalisation performance of the network and allow it a greater dimensionality for its representation of the data we choose to set the input and output size to 384. The input to the network is zero-padded to this length, and the output and latent space are concatenated with zero-padding in between. This width was also determined empirically, as it was found that the model had difficulties converging for the minimum width of 210. This problem was greatly mitigated by increasing the width, a technique that was found in the examples of [Ardizzone et al. \(2018\)](#), so no other values were trialled. The deep neural networks used inside the affine coupling layers have an input and output of length 192, but this is increased to 384 for the inner layers, to further increase their representational ability.

The RADYNVERSION model presented here considers each set of atmospheric parameters and observables as instantaneous quantities. As we have discussed, it is usually necessary

## 7.2 Training Data

to consider time-dependent populations in flares, and clearly we do not do so here. The atomic populations are not considered directly with this model, thus we are interested in determining whether there is sufficient information present in this reduced description of the atmosphere (under the inherent assumption in the training set that all snapshots originate from the same starting atmosphere) to reproduce the emergent line profiles and whether an ANN can then learn to decode this information.

## 7.2 Training Data

To train this model we use 81 RADYN simulations computed by the F-CHROMA project<sup>3</sup>. These models all start from a variant of the VAL3C quiet sun atmosphere (Vernazza et al. 1981), slightly modified to remain stable in RADYN. All models are heated by a symmetric triangular electron beam pulse, modelled using the Fokker-Planck module (with an initial power law distribution of electron energies), of 20 s duration, with a peak at 10 s. The total beam deposition varies between  $3 \times 10^{10}$  erg cm<sup>-2</sup> and  $1 \times 10^{12}$  erg cm<sup>-2</sup>, the low-energy cut-off is one of {10, 15, 20, 25} keV, and the spectral index of the electron energy distribution is one of {3, 4, 5, 6, 7, 8}. All of these simulations last for 50 s, with data saved every 0.1 s.

Not all of these simulations converged due to certain parameter combinations. In particular some simulations with lower values for the low-energy cut-off, higher spectral indices, and high total energy deposition were not present in the grid of models. From the 81 simulations we then have 40,500 individual timesteps, of which we separate 20% for validation purposes. The atmospheric parameters and line profiles are mapped onto their fixed grids and the parameters are prepared as discussed previously. Our height stratification is chosen to primarily sample the chromosphere, and places 45 linearly spaced points below 3.5 Mm, with a constant spacing of 79.2 km. The remaining 5 points are then exponentially spread through the corona from 3.5 Mm to 10 Mm.

## 7.3 Training Method

Our training method is based on the one presented in Ardizzone et al. (2018) and the network is constructed using their framework, FrEIA<sup>4</sup>. The INN is trained in both directions to ensure the conditioning of both the forward and inverse problems. The model is

---

<sup>3</sup><https://star.pst.qub.ac.uk/wiki/public/solarmodels/start.html>

<sup>4</sup><https://github.com/VLL-HD/FrEIA>

### 7.3 Training Method

trained using minibatching with the same minibatch of the training set being used in both directions. Both training directions are constrained by two loss functions, and a linear combination of these serves as our final loss to minimise. The forward direction (from atmospheric parameters to line profiles) uses an L2 loss ( $\|y - y_{\text{true}}\|_2^2$ , where  $y$  indicates the output of the network and  $y_{\text{true}}$  the expected output) on the output vector of line profiles and zero padding. This latent space is constrained by a Maximum Mean Discrepancy (MMD) loss. The MMD is a loss that compares distributions from finite samples, and is computed between batches of  $[y, z]$  and  $[y_{\text{true}}, \mathcal{N}(0, J_z)]$ . This is discussed in depth, along with implementation details in Section 7.4. During the forward process the MMD loss is used to ensure that the network learns to map the true latent space to our chosen form for it (the multivariate unit Gaussian distribution). A traditional regression loss cannot be applied here, as we would have to assign fixed samples from the latent space for each timestep in the training set, which cannot be done without understanding the true latent space. Its aim is instead to condition the form of the distribution produced in the latent space. To this end, whilst  $y$  is included in the MMD loss terms (as this is an important component of the output), the gradients on  $y$  due to the MMD loss are ignored, so as not to affect the training of the forward model i.e. whilst the elements of  $y$  are considered when computing the value of the MMD loss, the MMD can only affect the weights determining  $z$ . The convergence of both the L2 and MMD losses ensures that samples of  $z$  are correctly independent of  $y$  as they must not contain copies of the same information (i.e. they must be independent) for the reverse process to work correctly.

The inverse process is trained similarly, with an additional two losses. An L2 loss is used for  $x$  and the zero-padding to ensure the expected atmosphere parameters are produced (and that the padding remain 0), and an MMD loss ensuring the correct distribution of  $x$  for random latent samples.

Both of the forward and backwards losses are linearly combined to produce a set of weighted gradients used to update the network. We define three hyperparameter weights for this purpose  $w_{\text{pred}}$ ,  $w_{\text{latent}}$ , and  $w_{\text{rev}}$ . These are combined to produce the losses

$$\text{loss}_f = w_{\text{pred}}L2_f + w_{\text{latent}}\text{MMD}_f, \quad (7.5)$$

$$\text{loss}_b = 0.5w_{\text{pred}}L2_b + \xi(n)w_{\text{rev}}\text{MMD}_b, \quad (7.6)$$

where  $f$  and  $b$  represent the forward and backward terms respectively, and  $\xi$  is a term that gradually increases towards unity over the course of a fade-in period. It is parametrised

### 7.3 Training Method

as

$$\xi(n) = \left( \min \left( \frac{n}{0.4N_{\text{fade}}}, 1 \right) \right)^3, \quad (7.7)$$

where  $n$  is the current epoch, and  $N_{\text{fade}}$  is the fade-in period. This term serves to slowly increase the effect of the MMD loss, as this was otherwise found to hinder the training of both the forward and inverse problems by dominating  $\text{loss}_b$  and guiding the system away from a solution that minimised both terms. To ensure that the output in the zero-padded sections remains close to zero we also use  $1 - \xi$  to initialise these with a small amount of random noise, decaying over this same period. This increases the activation of these neurons early on, forcing the network to learn that these must be adjusted towards 0 for all inputs. Through a manual hyperparameter search we found that for a fade-in period of 800 epochs, a good choice for the loss weights was  $w_{\text{pred}} = 4000$ ,  $w_{\text{latent}} = 900$ , and  $w_{\text{rev}} = 1000$ . After the fade-in period the network was trained in blocks of 400 epochs, increasing the value of  $w_{\text{pred}}$  by 1000 for each of these blocks to fine-tune the line profiles and atmospheric parameters which were otherwise a little noisy. From 4,800 epochs to 12,000 epochs the network was trained in blocks of 600 epochs, with the value of  $w_{\text{pred}}$  again being increased by 1000. These weights were all tuned empirically, and others were also found to yield good convergence, however we found it important that the L2 weight be a factor of 2 or more larger than the MMD weights or the forward process would not reliably converge.

The update of the weights inside the neural networks is computed from the gradients of the linearly combined losses using the Adam optimisation algorithm (Kingma & Ba 2014) with hyperparameters  $\beta_1 = \beta_2 = 0.8$  and  $\epsilon = 1 \times 10^{-6}$ . The  $\beta_i$  terms control the decay rate for momentum of the first- and second-moment estimates of the gradients and  $\epsilon$  simply prevents division by zero. These gradients are limited to a range of  $\pm 15$  to help further mitigate problems with exploding gradients. This does not affect the final solution as the gradients will become small as we approach a minimum. The learning rate was initialised to  $1.5 \times 10^{-3}$  and decays by a factor of  $0.004^{1/1333}$  every 12 epochs. Each minibatch contained 500 different samples and 20 minibatches were used per learning epoch. The final model was selected based on its L2 performance for the forward and backwards results on the validation set. In this case the best performing model was the one saved after 11,400 epochs of training.

## 7.4 Maximum Mean Discrepancy

The content of this section draws primarily from *Sriperumbudur et al. (2009)*, *Gretton et al. (2012)*, *Muandet et al. (2017)*, and Gretton's lecture content, currently available online<sup>5</sup>.

The maximum mean discrepancy (MMD) is a statistic for comparing two distributions, based on samples drawn from these, and computes the difference in expectations over functions in the unit ball of a reproducing kernel Hilbert space (RKHS). It represents an estimate of the distance between the mean features of the sampled distributions in the space defined by a chosen kernel i.e. the greater similarity between the two sampled distributions the smaller the value of the MMD.

A Hilbert space is a space in which an inner product is defined that is linear, symmetric, and where the inner product of an element with itself is positive definite. A norm can be defined from this inner product  $\|f\| = \sqrt{\langle f, f \rangle}$ . These spaces must also be Cauchy complete, implying that every Cauchy sequence (convergent sequence), must converge to a point in the space.

An RKHS  $\mathcal{H}$  is then a Hilbert space of functions  $f : \mathcal{X} \rightarrow \mathbb{R}$  for which the evaluation functional  $\delta_x : f \mapsto f(x)$  is bounded and continuous<sup>6</sup>. The RKHS has the property that two functions that are close in norm in  $\mathcal{H}$  are then pointwise close when evaluated anywhere over  $\mathcal{X}$ . The kernel associated with this RKHS is a positive definite kernel  $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  if there exists a map  $\phi : \mathcal{X} \rightarrow \mathcal{H}$  such that  $\forall x, y \in \mathcal{X}$

$$k(x, y) = \langle \phi(x), \phi(y) \rangle_{\mathcal{H}}. \quad (7.8)$$

The map  $\phi$  is termed the feature map, and  $\mathcal{H}$  is known as the feature space.

Let us define two probability distributions  $P$  and  $Q$  and draw observations  $X$  and  $Y$  in an independent and identically distributed fashion from  $P$  and  $Q$  respectively. The MMD is then defined by

$$\begin{aligned} \text{MMD}^2 &= \|\mu_P - \mu_Q\|_{\mathcal{H}}^2 \\ &= \langle \mu_P, \mu_P \rangle_{\mathcal{H}} + \langle \mu_Q, \mu_Q \rangle_{\mathcal{H}} - 2\langle \mu_P, \mu_Q \rangle_{\mathcal{H}}, \end{aligned} \quad (7.9)$$

where  $\mu_A$  represents the mean embedding of the distribution  $A$  in the feature space, which is the expectation vector of the features of  $\mathcal{H}$  evaluated for this distribution. The

<sup>5</sup><http://www.gatsby.ucl.ac.uk/~gretton/teaching.html>

<sup>6</sup>This can be generalised to functions  $f : \mathcal{X} \rightarrow \mathbb{C}$ .

#### 7.4 Maximum Mean Discrepancy

kernel associated with  $\mathcal{H}$  is *characteristic* if the feature map is injective, and in this case the MMD is zero iff  $P = Q$ . It was shown by [Sriperumbudur et al. \(2009\)](#) that all measurable and bounded strictly positive definite kernels are characteristic, and we therefore limit ourselves to this class.

$\mu_P$  can now be written in terms of the features of  $\mathcal{H}$

$$\mu_P = [\dots \mathbb{E}_P[\phi_i(X)] \dots], \quad (7.10)$$

where  $\mathbb{E}_P$  denotes the expected value of its argument with respect to  $P$ , and  $\phi_i$  is the  $i$ th feature of the feature map (which may contain infinitely many features). We can then write

$$\langle \mu_P, \mu_P \rangle_{\mathcal{H}} = \langle \mathbb{E}_P[k(\cdot, X)], \mathbb{E}_P[k(\cdot, X')] \rangle_{\mathcal{H}} = \mathbb{E}_P[k(X, X')], \quad (7.11)$$

where  $X'$  is an independently drawn copy of  $X$  from  $P$ , and  $k(\cdot, X)$  refers to the function, and not its evaluation at a particular point. Thus  $k(\cdot, X)$  is the feature map  $\phi(X)$ . We define  $Y'$  analogously for  $Y$  and  $Q$  to then write (7.9) as

$$\text{MMD}^2 = \|\mu_P - \mu_Q\|^2 = \mathbb{E}_P[k(X, X')] + \mathbb{E}_Q[k(Y, Y')] - 2\mathbb{E}_{P,Q}[k(X, Y)]. \quad (7.12)$$

Here the first two terms compare the distributions for internal similarity, whereas the last compares the intra-distribution similarity.

For length  $n$  finite observations  $X$  and  $Y$  we can expand this to provide an unbiased sample estimate

$$\widehat{\text{MMD}}_{\text{u}}^2 = \frac{1}{n(n-1)} \sum_{i \neq j} k(x_i, x_j) + \frac{1}{n(n-1)} \sum_{i \neq j} k(y_i, y_j) - \frac{2}{n^2} \sum_{i,j} k(x_i, y_j). \quad (7.13)$$

As this statistic is unbiased and computed from a finite sample size, it can be negative when  $P$  and  $Q$  are similar distributions, despite being the definition of  $\text{MMD}^2$ . We therefore employ a biased estimate of the MMD which remains positive in all scenarios

$$\widehat{\text{MMD}}_{\text{b}}^2 = \frac{1}{n^2} \sum_{i,j} (k(x_i, x_j) + k(y_i, y_j) - 2k(x_i, y_j)). \quad (7.14)$$

This biased form of the MMD is also more efficient to compute using the vectorised tensor operations present in machine learning frameworks like PyTorch.

The choice of kernel used in the MMD will then determine how this statistic is able to distinguish between distributions. We choose to employ the inverse multi-quadric (IMQ)

## 7.5 Validation

kernel as used by Tolstikhin et al. (2017) and Ardizzone et al. (2018)

$$k_{\alpha}(x, y) = \frac{\alpha^2}{\alpha^2 + \|x - y\|_2^2}. \quad (7.15)$$

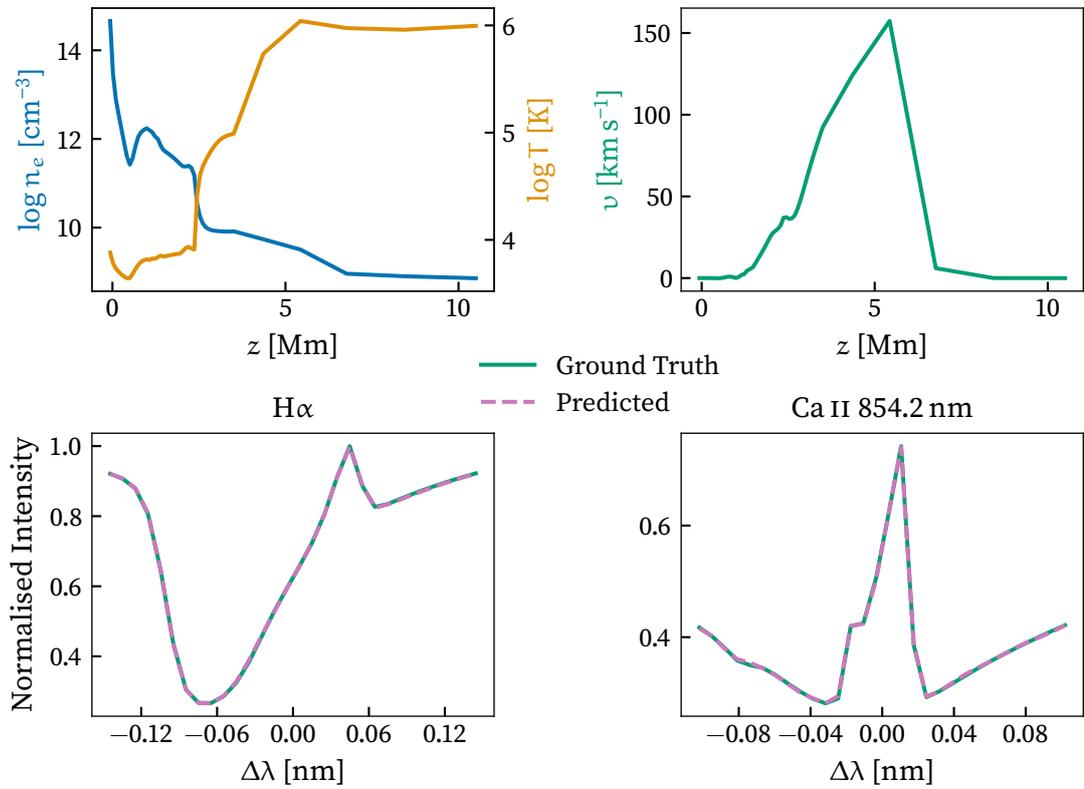
This kernel meets the strictly positive definite criterion that is necessary for a characteristic kernel of an RKHS, but contains a free parameter  $\alpha$ . This equation therefore describes a family of kernels, which will have different performance depending on the data and the choice of  $\alpha$ . We were unable to empirically determine a single value for  $\alpha$  for which the MMD could reliably distinguish between the two distributions well enough to optimise over, however it was found that the value of the MMD for different values of  $\alpha$ , whilst retaining fixed  $X$  and  $Y$ , peaked for a particular value of  $\alpha$ . The kernel of the family with this value of  $\alpha$  is able to better resolve the differences between our fixed samples and is therefore best suited (of the tested values) to guiding the optimisation of the weights associated with these terms. This approach was used to refine the MMD losses by updating the  $\alpha$  parameter of the kernel to the value for which the MMD was maximal (given the fixed input samples from this epoch) every five epochs of training. Modifying the MMD loss in this way ensures that it is most sensitive to the scale of the currently differing features between our distribution samples.

Minimising the forward MMD loss then ensures the independence of data stored in the latent space and that the latent space takes a normal unit distribution form, whilst minimising the backwards MMD ensures that the distribution of atmospheres generated with different draws of the latent space is a plausible distribution when compared against the training set.

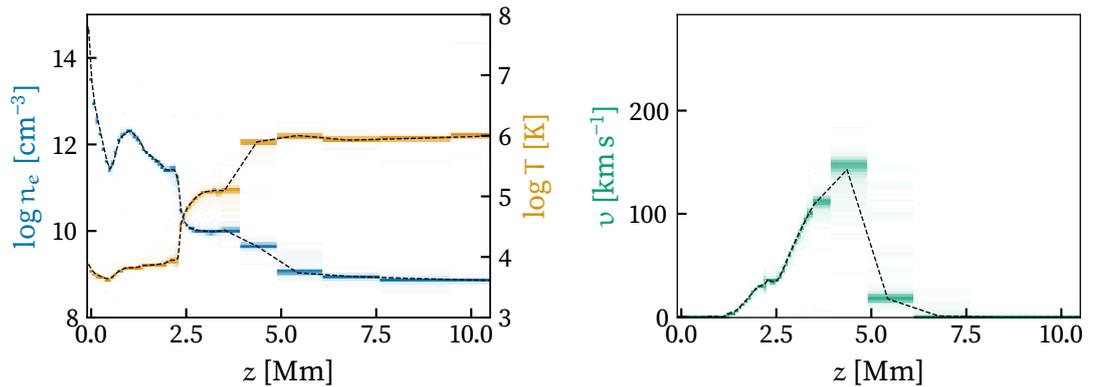
## 7.5 Validation

During and after training, the RADYNVERSION model is validated against the unseen validation data separated from the training set. An example taken from the validation set showing the forward process is presented in Fig. 7.2. The upper row shows the atmosphere input into the model whilst the lower row shows the expected output (solid lines) and the output of the forward process (dashed lines). The mean squared error in the scaled intensity at each wavelength point for the validation set is  $5.74 \times 10^{-5}$  showing that the model can very accurately predict the scaled line profiles based on the atmospheric parameters for unseen atmospheres.

## 7.5 Validation



**Figure 7.2:** Validation of the RADYNVERSION forwards process on unseen data. The upper row shows the atmospheric input, and the lower row the expected output along with the prediction from the network.



**Figure 7.3:** Validation of the RADYNVERSION reverse process from the ground truth line profiles shown in Fig. 7.2. The two-dimensional histograms show the probability density of the solution in each altitude node, and the black dashed lines the expected solution.

## 7.6 Proof of Concept Results

Due to the inherent non-uniqueness and presence of possible degeneracies it is harder to numerically evaluate the performance of the model's reverse process. For each pair of line profiles we draw a large number of samples from the latent space and infer the predicted atmosphere from each latent space draw combined with the pair of line profiles. This gives a large number of possible sets of atmospheric parameters which we plot on a two dimensional histogram to show the probability density of the solution in each altitude node. This method then allows us to gauge possible degeneracies and establish the relative probabilities of different solutions. An example validation inversion is shown in Fig. 7.3, using the ground truth output line profiles shown in the forward validation process as input. These histograms were generated from 10 000 latent space draws, and their values were gamma corrected (with  $\gamma = 0.2$ ) to enhance the visibility of less probable solutions. The overplotted black dashed lines show the ground truth solution for each of the atmospheric parameters, and overlap the peak density of the histogram extremely well. The histogram remains very narrow in the lower atmosphere, but starts to expand above 3 Mm, where the solution is poorly constrained by the chromospheric lines in use here. This effect is most visible on the velocity plot, but despite the increase in uncertainty, the solutions remain accurate on the validation set due to the model's conditioning to the RADYN training data.

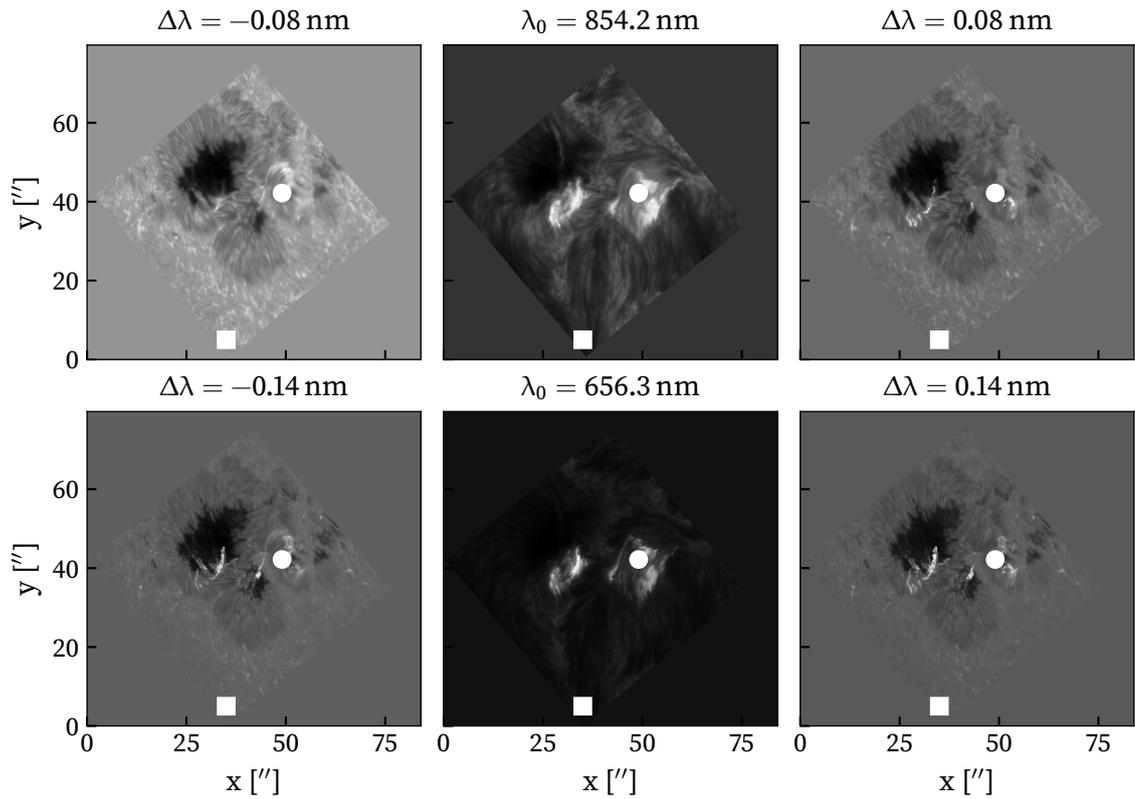
## 7.6 Proof of Concept Results

*The proof of concept application of the RADYNVERSION model was primarily undertaken by my co-authors. I will briefly summarise the primary results in this section, but the methodology is more fully described in the associated paper.*

As a proof of concept the RADYNVERSION model was applied to the two-ribbon M1.1 flare SOL 20140906T17:09 observed with the CRISP instrument in H $\alpha$  and Ca II 854.2 nm. The preparation of this data is briefly discussed in Sec. 3.1.1.

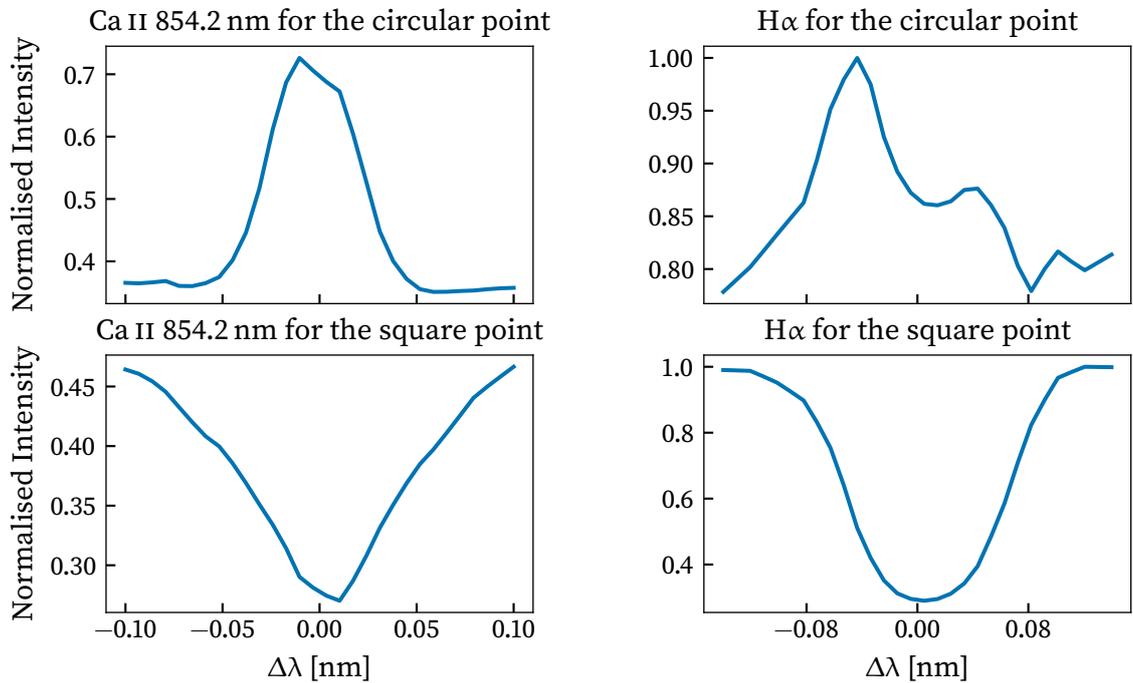
Due to the formation heights of these spectral lines, our interest is primarily focused on the region below  $\sim 2$  Mm. Fig. 7.4 shows CRISP images in the blue wing, line core, and red wing for both of the spectral lines respectively. These images were taken just after flare onset. Two pixels are marked, a circle on the flare ribbon, and a square far from the flare, in a much quieter region. The spectral line profiles from these two pixels are shown in Fig. 7.5. For the circular point both lines are strongly in emission whereas for the square point they are broad absorption lines. These lines from these pixels have then been

7.6 Proof of Concept Results



**Figure 7.4:** Observations of the the M1.1 flare from AR 12157 on 2014-09-06 just after flare onset. The upper row shows images in the Ca II 854.2 nm band, and the lower row shows equivalent images from the H $\alpha$  band. These two inverted pixels are marked by the square (off-ribbon) and circle (on-ribbon). Wavelengths in this figure use their values in air.

## 7.6 Proof of Concept Results

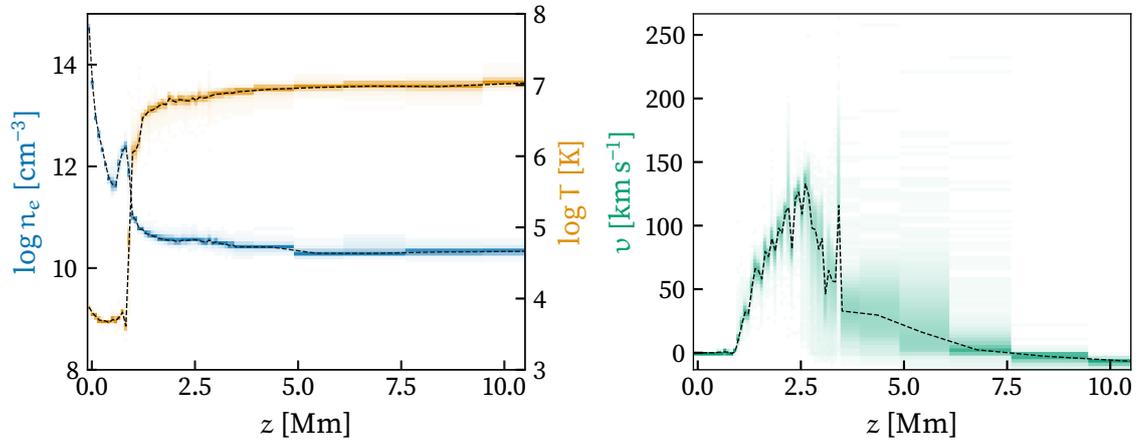


**Figure 7.5:** The spectral line profiles of Ca II 854.2 nm and H $\alpha$  for the two pixels marked in Fig. 7.4.

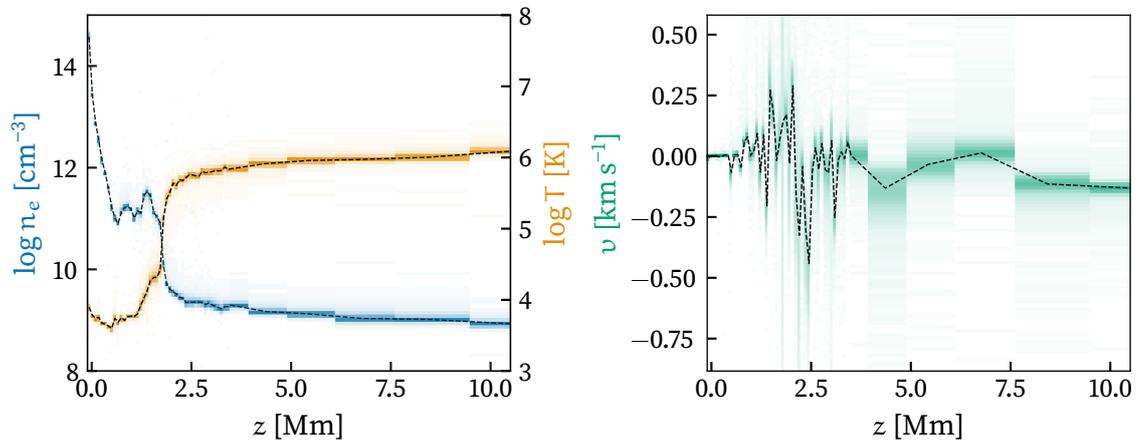
inverted using RADYNVERSION, with 20 000 latent space draws each. The results of these inversions are shown in Figs 7.6 and 7.7 for the circular and square points respectively. These figures present equivalent information to the lower panels of Fig. 7.3, although the dashed black lines now show the median of the histogram at each atmospheric point as there is no ground truth solution available. Performing these inversions with 20 000 draws each takes  $\sim 893$  ms on an NVIDIA GTX 1050 Ti GPU installed in *hephaistos*.

The following is a brief summary of the analysis undertaken by my co-authors. The inverted pixels were found to be consistent with previous analyses. For example, forward modelling by Kuridze et al. (2015) suggests that the H $\alpha$  line profile forms below 1.2 Mm, with the core forming towards the top of this region, and the line wings forming below 0.95 Mm. In this observation the H $\alpha$  line profile is found to be asymmetric in favour of the red wing, and from the inversions we see that there is a slight upflow in the region where the wings are formed. This is likely due to chromospheric evaporation in this region causing an increase in the opacity of the blue wing, leading in turn to more intensity in the red wing.

## 7.6 Proof of Concept Results



**Figure 7.6:** Inversion results for the on-ribbon pixel. The histograms represent the probability density for the solution in each altitude node, and the median solution is shown with the dashed black lines.



**Figure 7.7:** Inversion results for the off-ribbon pixel. The histograms represent the probability density for the solution in each altitude node, and the median solution is shown with the dashed black lines.

## 7.7 Discussion

For the off-ribbon pixel we note that both line profiles are very broad. The inverted velocity field contains a significant number of small magnitude oscillations which we believe are due to RADYN's conservative assumption of a  $2 \text{ km s}^{-1}$  microturbulent velocity throughout the atmosphere, as this is insufficient to produce line profiles as broad as those observed here. If the network has learned that bulk plasma flows can shift the position of the line core (i.e. the majority of the opacity in a line), then it is reasonable to suggest that these represent our model's attempt to broaden the line profiles. Recent studies have shown that significantly higher microturbulent velocities are needed to explain the non-thermal broadenings observed in chromospheric plages ( $6\text{--}7 \text{ km s}^{-1}$ , Carlsson et al. 2015). The other atmospheric parameters vary more smoothly and do not reach values as large as those found for the on-ribbon pixel, which is consistent with the much quieter atmosphere expected at this location.

## 7.7 Discussion

The RADYNVERSION invertible neural network presented in this chapter represents a novel method for investigating the atmospheric properties of observed events and lifts many of the restrictions that previously made inversions of flaring line profiles infeasible. This is achieved by the union of machine learning and a large body of RHD simulations. The implementation is currently a proof of concept but shows much promise, agreeing with previous investigations, and can easily be extended to other spectral lines and atmospheric parameters. With the addition of more advanced forward models this technique could be applied to inference of the chromospheric magnetic field from full Stokes observations. Inversions using this method are fast and robust; the effort of “exploring” the latent space that is common in regression based codes has effectively been replaced with an up-front cost in the training process. Once the model is trained, taking multiple draws from the latent space replicates the exploration, but extremely rapidly, due to the previous training effort which is then shared between every application of the model (contrary to the traditional 1.5D inversion technique, where every column is treated independently, and a lot of this work is replicated). These latent draws additionally serve to provide an estimate of the uncertainty on the inferred parameters. Potential future extensions to this model include incorporating the concept of time-dependence in a robust fashion so as to condition the solution from previous observations, and increasing the number of atmospheric parameters inferred.

The RADYNVERSION approach has the potential to provide new insight into the structure

## 7.7 Discussion

of the flaring chromosphere and allow investigation of complete observation fields of view in a timely fashion, allowing researchers to leverage the full capabilities of next-generation solar telescopes such as DKIST.

# 8

## Concluding Remarks

In this thesis we have investigated the formation and inversion of chromospheric optical spectral lines, in particular  $H\alpha$  and Ca II 854.2 nm. These lines form outside of local thermodynamic equilibrium (LTE) and can only be synthesised through a detailed treatment of the atmospheric radiation field. To facilitate the modelling of these spectral lines, we have developed the *Lightweaver* framework: a modular radiative transfer Python package capable of handling both plane-parallel and two-dimensional geometries. The purpose of a framework such as this is to empower researchers with the ability to easily create custom tools for the radiative transfer problems they wish to simulate. The conceptual design of *Lightweaver*, along with a series of validation examples, were presented in Chap. 4, building on the radiative transfer theory presented in Chap. 2. The extension of *Lightweaver* to support two-dimensional atmospheres was presented in Chap 6. The development of *Lightweaver* has enabled much of the research presented in this thesis, thanks to its flexibility, and we hope that other researchers will be able to make use of it<sup>1</sup>.

We have presented the application of *Lightweaver* to the synthesis of spectral lines with time-dependent populations in both plane-parallel flaring simulations and the irradiation of a slab of quiet Sun atmosphere by an adjacent flare model. In Chap. 5 we used *Lightweaver* to investigate some of the assumptions present in flare models produced by the most commonly used state-of-the-art radiation hydrodynamic code RADYN. We performed in-depth investigations of the effects of the hydrogen Lyman lines on the Ca II populations (and thus emergent line profiles), and discussed whether a full time-dependent treatment of the Ca II populations is needed. We also presented some of the difficulties encountered when trying to treat flaring models with time-dependence and partial frequency redis-

---

<sup>1</sup>*Lightweaver* (Osborne & Milić 2021) is developed openly under the MIT license on GitHub (<https://github.com/Goobley/Lightweaver>), with archival on Zenodo (Osborne 2021a).

tribution simultaneously. The concept of a time-dependent response function was also introduced as a new tool for analysing RHD models.

The atmospheric evolution of two different RADYN simulations (with constant F9 and F10 beams) was used as the input for a *Lightweaver*-based tool. The synthetic spectra produced serve both as a validation of the time-dependent radiative transfer techniques implemented in *Lightweaver*, and also a validation of RADYN. Differences were found between the Ca II lines synthesised with RADYN and *Lightweaver*, and were found to result from photoionisation by the Lyman lines. RADYN's default includes the photoionising effects of the Lyman continuum on Ca II, but neglects the impact of the Lyman lines. The Lyman series contains some of the strongest spectral lines in the flaring solar spectrum, with extreme enhancements observed over their quiet Sun values. The additional flux produced by these lines is sufficient to provoke substantial changes in the synthesised Ca II line profiles, in particular, that of the Ca II 854.2 nm line. The effects of the Lyman lines on the shape and intensity of the Ca II line profiles also impact the net radiative losses from the atmosphere. In the simulations there was found to be a 10–15 % variation in the chromospheric radiative losses. This difference could plausibly change the atmospheric evolution of the simulation, leading to greater differences in the Ca II line profiles from models self-consistently taking these effects into account, and also possibly changing the observed line profiles from other species such as hydrogen.

We also investigated the necessity of performing a fully time-dependent treatment of the Ca II level populations, once again using *Lightweaver* to reprocess a RADYN simulation. It was found that for a significant majority of the simulation, there was little difference between line profiles computed with statistical equilibrium populations and those computed with a time-dependent treatment. The most significant differences occurred at the start of the simulation, as the atmosphere first reacts to the heating, but these effects are relatively minor. The possibility of treating this species in statistical equilibrium may provide future opportunities for further optimisation.

We implemented a method for including the effects of partial frequency redistribution into the *Lightweaver*-based tool developed in Chap. 5. This was found to be a difficult problem, often suffering from non-convergence, but we were able to fully reprocess the previously discussed F9 simulation. The Doppler-like line profile approximations implemented in RADYN for the Ly $\alpha$  and Ly $\beta$  proved to be relatively accurate at most points in the simulation. Much larger differences were found for the calcium resonance lines, with significant differences in the radiative losses due to these lines. It therefore seems essential to develop approximate treatments for the Ca II H & K spectral lines.

The concept of time-dependent response functions was also introduced in this chapter: these generalise the statistical equilibrium response functions for atomic level populations that cannot be treated in a time-independent fashion. These also represent a powerful tool for disambiguating the atmospheric response to *different* thermodynamic parameters, whilst also considering the natural settling time of the populations towards the statistical equilibrium solution. For species such as hydrogen, which are found to have a long settling time, a time-dependent treatment is necessary to obtain the instantaneous atmospheric response to a parameter, especially when including the effects on the electron density due to charge conservation.

In Chap. 6, we investigated the radiative response and outgoing line profiles from a two-dimensional slab of quiet Sun illuminated by an adjacent column of flaring plasma. Despite the thermodynamic properties of the slab being fixed to the initial quiet Sun atmosphere (other than the electron density, which was allowed to vary to ensure charge conservation), significant enhancements of the H $\alpha$  and Ca II 854.2 nm lines were found 1 Mm and further from the flaring boundary. This has several implications. Firstly, any kind of traditional column-by-column inversion technique employed on regions adjacent to flaring ribbons is likely to be led astray, and infer an incorrect atmospheric structure. This is due to the atomic level populations being determined by a non-local *transverse* radiation field, rather than the plasma parameters in the column. Secondly, the enhancements produced in our simulations are far from uniform as a function of wavelength, being much more dramatic in the line core than in the wings, with no effect being seen in the continuum. The lack of continuum enhancement is primarily due to the temperature and mass density structure in the slab being fixed, whilst the continuum forms in approximately LTE conditions, and is therefore dependent on the local thermodynamic parameters. As such, calculations of filling factors may need to take this wavelength dependence into account. Finally, the enhancements observed in our models are on a scale already easily resolvable with modern ground-based solar telescopes, such as the SST, and are likely to become more important with the generational leap provided by DKIST. We also presented a simple comparison against SST/CRISP observations and found enhancements on the same order of magnitude as those produced by our simple model. When analysing the regions adjacent to flaring ribbons in optical spectral lines, it therefore will be necessary to take into account the effects of horizontal irradiation such as this.

RADYNVERSION, a novel machine learning inversion technique, was presented in Chap. 7. This model uses an invertible neural network to simultaneously learn both the forward (synthesis) and inverse problems of radiative transfer based on atmospheric snapshots produced with RADYN. Thanks to this training set, the model learns to synthesise the H $\alpha$

and Ca II 854.2 nm spectral lines from a geometric stratification of temperature, electron density, and velocity. It is the first non-LTE inversion technique not constrained by the assumptions of statistical equilibrium and hydrostatic equilibrium, rendering it much more applicable to flares than conventional approaches. RADYNVERSION learns the possible information lost in the forward process from its training set, and we are then able to sample this to produce posterior distributions for the stratified atmospheric parameters. A proof of concept analysis of two pixels observed with the SST/CRISP instrument was shown to be in accord with previous investigation of the same event involving forward modelling using RADYN by Kuridze et al. (2015). This model is also extremely performant once trained, taking  $\sim 10 \mu\text{s}$  per latent space draw on modest consumer computing hardware, as its training process “front-loads” a lot of the work that is undertaken repeatedly in a regression-based inversion model (although this can be reduced by the use of database initialisation techniques). Enhancing the performance of inversion techniques, through approaches such as RADYNVERSION, is essential to a detailed exploitation of the vast quantity of data produced by current observatories that will only be dwarfed by those arriving in the coming solar cycle.

## Future Directions

The *Lightweaver* framework is robust and has been proven production-ready by the applications presented in this thesis. Nevertheless, there are many enhancements that could yet be implemented. These include:

- The inclusion of more rapid iteration schemes such as forth-and-back implicit lambda iteration (Atanacković-Vukmanović et al. 1997; Kuzmanovska et al. 2017), or the hybrid scheme described in Avrett & Loeser (2008) for possibly improving the handling of strong lines that exhibit partial frequency redistribution effects.
- Full Stokes synthesis is currently supported in plane-parallel models, but only an unpolarised formal solver has been implemented for the two-dimensional case.
- *Lightweaver* could be extended to treat three-dimensional radiative transfer. The frontend of the framework is already designed to support this, but no formal solver has been implemented.
- To support the use of *Lightweaver* in two- and three-dimensional modelling, a domain-decomposition technique to split large simulations across clusters of ma-

chines is likely to be beneficial. An MPI-based implementation has already been tested for the 1.5D column-by-column situation, but full domain-decomposition will require modifications to the core of *Lightweaver* itself.

- The equation of state used in the front-end is written in pure Python, and is very slow, often taking longer than the full non-LTE calculations for a simple plane-parallel atmosphere on a parallel machine. This should be replaced with a more performant, and possibly more advanced, implementation.
- *Lightweaver* could serve as the base of an inversion package, and the technique used for computing the response functions would determine the scale of the modifications required. For a STiC (de la Cruz Rodríguez et al. 2019) style approach, no modifications would be needed, as the machinery for computing finite-difference response functions is already present. If a SNAPI (Milić & van Noort 2018), or DeSIRE (B. Ruiz Cobo et al. *in preparation*) style approach were instead taken, the implementation of a technique for computing the necessary analytic response functions would be needed.
- Most RT codes, including *Lightweaver*, use fixed wavelength quadratures for each spectral line, and this requires pessimistically determining the minimum resolution needed to evaluate the necessary integrals. It could be highly beneficial to introduce an adaptive wavelength quadrature that can estimate the error in these integrals, and refine if necessary. An initial implementation of this would likely use a form of step-doubling techniques.
- Graphical processing units (GPUs) are becoming ever more ubiquitous and powerful, and can easily be adapted to radiative transfer calculations. A well-optimised GPU implementation of the routines needed for non-LTE radiative transfer can likely provide an order of magnitude increase in performance at similar hardware cost.
- The scattering integral needed to evaluate the effects of partial frequency redistribution could likely be accurately approximated by a neural network. This could provide significant performance improvements in angle-dependent PRD calculations where the evaluation of the scattering integral is highly computationally intensive.

This list is far from exhaustive, but it should be clear that while *Lightweaver* is a powerful and flexible package, there are many interesting directions to explore.

Another interesting application of *Lightweaver* would be to incorporate it into a radiation hydrodynamic modelling package. Its advanced radiative transfer could then be used in

self-consistent field-aligned radiation hydrodynamic modelling. Of the currently available codes, HYDRAD is the most suited to this treatment, in this author’s opinion. It is available on GitHub<sup>2</sup> under the MIT license, and consists of a relatively small body of C++ that could be bound to Python in a similar way to *Lightweaver*, with Python controlling the flow of data between the two. This would allow investigation of the magnitude of the Ca II photoionisation effects discussed in Chap. 5, whilst self-consistently considering the modified energy balance. Of course, this effect can also be incorporated into RADYN or FLARIX by including the radiation field of the Lyman lines in the calculation of the Ca II continua. This would likely need to be done under the assumption that the Ca II continua are not a primary source of opacity at these wavelengths, as this would require the treatment of the Lyman lines and Ca II continua to be coupled. In the modelling we have undertaken, this appears to be a safe assumption. Similarly to Ca II, Mg II may be photoionised by the hydrogen Lyman lines. We feel that the effect is likely to be of lesser magnitude, as the Mg II resonance continuum edge is situated at 82.46 nm, and the Lyman lines could therefore only affect the subordinate continua of this ion. Nevertheless, a similar investigation should be conducted to determine the scale of these effects.

We found that statistical equilibrium was a good approximation for the calcium level populations at most points in our flaring models. An investigation of whether statistical equilibrium can also be employed for hydrogen in flare models is also needed. It is clear that the non-equilibrium ionisation state of hydrogen needs to be known, but whether its level population distribution can be treated in statistical equilibrium whilst taking this ionisation into account remains to be seen in flaring models. If the primary species needed to compute chromospheric radiative losses in flare models could be accurately treated in some form of statistical equilibrium this would allow neural networks to be much more easily applied to this problem. Similarly to RADYNVERSION, this has the potential to greatly accelerate this style of simulation, even if a slight accuracy trade-off occurs.

The two-dimensional slab model we presented is intended to serve as a “first-order” approximation of the situation, and there are clearly many effects that were not included here. A simple extension of this model would include a method by which the temperature of the plasma in the slab could change based on the radiation absorbed. Such a model should also likely consider the effects of heat-conduction, even if the magnetohydrodynamics of such a situation are not considered. In our opinion, allowing the plasma to be heated by the radiation from the neighbouring flare is likely to increase the magnitude

---

<sup>2</sup><https://github.com/rice-solar-physics/HYDRAD>

of the effects seen in Chap. 6, and it is possible that this heating could produce continuum enhancements and affect spectral lines forming significantly deeper in the atmosphere.

Another development of the two-dimensional flare model would be to investigate the limitations of the plane-parallel radiative transfer model assumed by the field-aligned radiation hydrodynamic codes. The plane-parallel treatment of the radiative transfer equation considers that the stacked slabs of homogeneous atmosphere are infinitely wide, whereas flux tubes appear to be quite narrow, especially in the chromosphere. It is plausible that the radiative losses and spectral lines produced by these models will differ significantly when embedded in a plasma with non-uniform opacity, and when the compact heated regions are not modelled as having infinite transverse extent. This could be modelled in a similar way to the work undertaken in Chap. 6, but instead embedding the RADYN simulation in the centre of the slab, rather than using it as a boundary condition. To model this, assumptions would need to be made regarding the diameter of the flux tube, although a series of models could be performed with different diameters, constrained by observations and radiative magnetohydrodynamic modelling. The size and intensity profile of any core-halo effects produced by these models could be used with observations to attempt to constrain the size of flare kernels at different depths in the atmosphere, using different spectral lines. Of course, this improvement could also be combined with the prior, allowing the temperature in the plasma surrounding the flare model to change, yielding a more advanced treatment of the physics with an emphasis on the radiation, rather than the dynamics, of the situation.

The techniques employed in the RADYNVERSION model are not unique to the spectral lines we presented. Indeed, a similar model can support any line formation problem, given the correct training set. Trained with the spectral lines we have presented here, RADYNVERSION is able to infer the atmospheric properties throughout the majority of the flaring chromosphere. These lines will lose sensitivity at higher temperatures, especially going into the transition region. As such, other lines such as Mg II h & k should be incorporated into this model, although this will create calibration and alignment difficulties if added to the current H $\alpha$  and Ca II 854.2 nm model due to the lack of an instrument which can observe all of these. Ca II H & K can be observed from the same telescope as H $\alpha$  and Ca II 854.2 nm (e.g. SST with CRISP and CHROMIS) so may represent an interesting addition to the RADYNVERSION model, but such a model would need to be trained to determine how much information was added by this additional data. All of Ca II H & K and Mg II h & k need to be modelled taking into account the effects of partial frequency redistribution, so the RADYN output could not be used to directly train such a model. Additionally, the Ca II photoionisation effects discussed in Chap. 5

had significant effects on the outgoing line profiles and should be taken into account in an updated RADYNVERSION model. This would require generating a new grid of models using an updated version of RADYN or reprocessing pre-existing models using *Lightweaver*, although this is dependent on the necessary output having been generated and preserved.

The quality of inference obtained from models such as RADYNVERSION is entirely dependent on the quality of the training data: a larger and more varied training set, as well as models that are trained for multiple different viewing angles will all be key to the development of RADYNVERSION into a dependable, widely applicable, inversion tool.

## Closing Remarks

The complexity of spectral line formation in the chromosphere means that there is always “more to do” to glean greater understanding of the atmospheric conditions at play. In this thesis we have primarily focused on modelling the formation of  $H\alpha$  and Ca II 854.2 nm, but the techniques presented are much more widely applicable. Regular high-resolution observations of these spectral lines with the SST, as well as the imminent arrival of the DKIST, drives us to consider the formation and analysis of these lines on more compact spatial scales where the non-uniformity of the solar atmosphere becomes increasingly important. This thesis is intended to be a step along this path, focusing more on the radiative treatment of these events, rather than the dynamic. All of the treatments presented here can be extended to support more complex configurations with additional physics, but we have shown that the effects of the hydrogen Lyman lines on Ca II cannot be neglected, that the radiation produced by moderate flare simulations is capable of provoking significant changes in the atomic level populations and radiative output of an adjacent slab of quiet Sun, and that the combination of modelling and machine learning can provide techniques to render the inversion of flaring spectral lines outside of the assumptions of statistical equilibrium and hydrostatic equilibrium tractable.

# Bibliography

- Abadi, M., Agarwal, A., Barham, P., et al. 2016. <https://arxiv.org/abs/1603.04467> 67
- Abbett, W. P., & Hawley, S. L. 1999, *The Astrophysical Journal*, 521, 906, doi: [10.1086/307576](https://doi.org/10.1086/307576) 12
- Acton, L. W., Leibacher, J. W., Canfield, R. C., et al. 1982, *The Astrophysical Journal*, 263, 409, doi: [10.1086/160513](https://doi.org/10.1086/160513) 49
- Allred, J. C., Alaoui, M., Kowalski, A. F., & Kerr, G. S. 2020, *The Astrophysical Journal*, 902, 16, doi: [10.3847/1538-4357/abb239](https://doi.org/10.3847/1538-4357/abb239) 14
- Allred, J. C., Hawley, S. L., Abbett, W. P., & Carlsson, M. 2005, *The Astrophysical Journal*, 630, 573, doi: [10.1086/431751](https://doi.org/10.1086/431751) 13
- Allred, J. C., Kowalski, A. F., & Carlsson, M. 2015, *The Astrophysical Journal*, 809, 104, doi: [10.1088/0004-637X/809/1/104](https://doi.org/10.1088/0004-637X/809/1/104) 12, 13
- Ardizzone, L., Kruse, J., Wirkert, S., et al. 2018, arXiv e-prints, arXiv:1808.04730. <https://arxiv.org/abs/1808.04730> 173, 174, 175, 176, 181
- Armijo, L. 1966, *Pacific Journal of Mathematics*, 16, 1, doi: [10.2140/pjm.1966.16.1](https://doi.org/10.2140/pjm.1966.16.1) 92
- Armstrong, J. A., & Fletcher, L. 2021, *Monthly Notices of the Royal Astronomical Society*, 501, 2647, doi: [10.1093/mnras/staa3742](https://doi.org/10.1093/mnras/staa3742) 51
- Arnaud, M., & Rothenflug, R. 1985, *Astronomy & Astrophysics*, 60, 425 21, 89
- Ascher, U. M., Ruuth, S. J., & Wetton, B. T. R. 1995, *SIAM Journal on Numerical Analysis*, 32, 797 45
- Asensio Ramos, A., & Díaz Baso, C. J. 2019, *Astronomy & Astrophysics*, 626, A102, doi: [10.1051/0004-6361/201935628](https://doi.org/10.1051/0004-6361/201935628) 171
- Asensio Ramos, A., Martínez González, M. J., & Rubiño-Martín, J. A. 2007, *Astronomy & Astrophysics*, 476, 959, doi: [10.1051/0004-6361:20078107](https://doi.org/10.1051/0004-6361:20078107) 62
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *Annual Review of Astronomy and Astrophysics*, 47, 481, doi: [10.1146/annurev.astro.46.060407.145222](https://doi.org/10.1146/annurev.astro.46.060407.145222) 90

## Bibliography

- Atanacković-Vukmanović, O., Crivellari, L., & Simonneau, E. 1997, *The Astrophysical Journal*, 487, 735, doi: [10.1086/304626](https://doi.org/10.1086/304626) 28, 192
- Auer, L. 1973, *The Astrophysical Journal*, 180, 469, doi: [10.1086/151976](https://doi.org/10.1086/151976) 28
- Auer, L. H., Fabiani Bendicho, P., & Trujillo Bueno, J. 1994, *Astronomy & Astrophysics*, 292, 599 35, 36
- Auer, L. H., & Heasley, J. N. 1976, *The Astrophysical Journal*, 205, 165, doi: [10.1086/154260](https://doi.org/10.1086/154260) 28
- Auer, L. H., & Mihalas, D. 1969, *The Astrophysical Journal*, 158, 641, doi: [10.1086/150226](https://doi.org/10.1086/150226) 28
- Auer, L. H., & Paletou, F. 1994, *Astronomy & Astrophysics*, 285, 675 19, 79, 84, 123
- Avrett, E. H., & Loeser, R. 2008, *The Astrophysical Journal Supplement Series*, 175, 229, doi: [10.1086/523671](https://doi.org/10.1086/523671) 111, 192
- Bates, D. R. 1952, *Monthly Notices of the Royal Astronomical Society*, 112, 40, doi: [10.1093/mnras/112.1.40](https://doi.org/10.1093/mnras/112.1.40) 75
- Battaglia, M., Fletcher, L., & Benz, A. O. 2009, *Astronomy & Astrophysics*, 498, 891, doi: [10.1051/0004-6361/200811196](https://doi.org/10.1051/0004-6361/200811196) 45
- Beckers, J. M., & Milkey, R. W. 1975, *Solar Physics*, 43, 289, doi: [10.1007/BF00152353](https://doi.org/10.1007/BF00152353) 59
- Behnel, S., Bradshaw, R., Citro, C., et al. 2011, *Computing in Science & Engineering*, 13, 31, doi: [10.1109/MCSE.2010.118](https://doi.org/10.1109/MCSE.2010.118) 71
- Bell, K. L. 1980, *Journal of Physics B: Atomic and Molecular Physics*, 13, 1859, doi: [10.1088/0022-3700/13/9/016](https://doi.org/10.1088/0022-3700/13/9/016) 75
- Benz, A. O. 2008, *Living Reviews in Solar Physics*, 5, 1, doi: [10.12942/lrsp-2008-1](https://doi.org/10.12942/lrsp-2008-1) 5
- Bian, N., Emslie, A. G., Horne, D., & Kontar, E. P. 2018, *The Astrophysical Journal*, 852, 127, doi: [10.3847/1538-4357/aa9f29](https://doi.org/10.3847/1538-4357/aa9f29) 46
- Bian, N. H., Watters, J. M., Kontar, E. P., & Emslie, A. G. 2016, *The Astrophysical Journal*, 833, 76, doi: [10.3847/1538-4357/833/1/76](https://doi.org/10.3847/1538-4357/833/1/76) 45
- Björgen, J. P., Leenaarts, J., Rempel, M., et al. 2019, *Astronomy & Astrophysics*, 631, A33, doi: [10.1051/0004-6361/201834919](https://doi.org/10.1051/0004-6361/201834919) 33, 50, 135
- Bradshaw, S. J., & Cargill, P. J. 2013, *The Astrophysical Journal*, 770, 12, doi: [10.1088/0004-637X/770/1/12](https://doi.org/10.1088/0004-637X/770/1/12) 12, 45
- Bradshaw, S. J., & Mason, H. E. 2003, *Astronomy & Astrophysics*, 401, 699, doi: [10.1051/0004-6361:20030089](https://doi.org/10.1051/0004-6361:20030089) 12, 45
- Braginskii, S. I. 1965, in *Reviews of Plasma Physics*, ed. M. A. Leontovich, Vol. 1 (New York: Consultants Bureau), 205 45

## Bibliography

- Brown, J. C. 1971, *Solar Physics*, 18, 489, doi: [10.1007/BF00149070](https://doi.org/10.1007/BF00149070) 6
- Brown, S. A., Fletcher, L., Kerr, G. S., et al. 2018, *The Astrophysical Journal*, 862, 59, doi: [10.3847/1538-4357/aacc29](https://doi.org/10.3847/1538-4357/aacc29) 93, 104, 107
- Burgess, A., & Chidichimo, M. C. 1983, *Monthly Notices of the Royal Astronomical Society*, 203, 1269 21
- Campbell, P. M. 1984, *Physical Review A*, 30, 365, doi: [10.1103/PhysRevA.30.365](https://doi.org/10.1103/PhysRevA.30.365) 45, 86
- Cannon, C. J. 1973, *The Astrophysical Journal*, 185, 621, doi: [10.1086/152442](https://doi.org/10.1086/152442) 26
- Capparelli, V., Zuccarello, F., Romano, P., et al. 2017, *The Astrophysical Journal*, 850, 36, doi: [10.3847/1538-4357/aa9187](https://doi.org/10.3847/1538-4357/aa9187) 49
- Carlsson, M. 1986, Uppsala Astronomical Observatory: Report No. 33: A Computer Program for Solving Multi-Level Non-LTE Radiative Transfer Problems in Moving or Static Atmospheres 12, 27, 28, 85
- Carlsson, M. 1992, in *Cool Stars, Stellar Systems, and the Sun*, Proceedings of the 7th Cambridge Workshop, ASP Conference Series, Vol. 26, ed. M. S. Giampapa & J. A. Bookbinder (San Francisco: Astronomical Society of the Pacific), 499 12, 27, 85
- Carlsson, M., De Pontieu, B., & Hansteen, V. H. 2019, *Annual Review of Astronomy and Astrophysics*, 57, 189, doi: [10.1146/annurev-astro-081817-052044](https://doi.org/10.1146/annurev-astro-081817-052044) 4, 5
- Carlsson, M., Hansteen, V. H., Gudiksen, B. V., Leenaarts, J., & De Pontieu, B. 2016, *Astronomy & Astrophysics*, 585, A4, doi: [10.1051/0004-6361/201527226](https://doi.org/10.1051/0004-6361/201527226) 13
- Carlsson, M., & Leenaarts, J. 2012, *Astronomy & Astrophysics*, 539, A39, doi: [10.1051/0004-6361/201118366](https://doi.org/10.1051/0004-6361/201118366) 12, 13
- Carlsson, M., Leenaarts, J., & De Pontieu, B. 2015, *The Astrophysical Journal Letters*, 809, L30. <https://arxiv.org/abs/1508.04888> 187
- Carlsson, M., & Stein, R. F. 1992, *The Astrophysical Journal*, 397, L59, doi: [10.1086/186544](https://doi.org/10.1086/186544) 12, 49
- . 1995, *The Astrophysical Journal*, 440, L29, doi: [10.1086/187753](https://doi.org/10.1086/187753) 12
- . 1997, *The Astrophysical Journal*, 481, 500, doi: [10.1086/304043](https://doi.org/10.1086/304043) 58
- Carlsson, M., & Stein, R. F. 1999, in *AIP Conference Proceedings*, Vol. 23 (AIP), 23–28, doi: [10.1063/1.58753](https://doi.org/10.1063/1.58753) 12
- . 2002, *The Astrophysical Journal*, 572, 626, doi: [10.1086/340293](https://doi.org/10.1086/340293) 88, 104, 112, 121, 158
- Carmichael, H. 1964, *NASA Special Publication*, 50, 451. <http://adsabs.harvard.edu/abs/1964NASSP..50..451C> 5
- Carrington, R. C. 1859, *Monthly Notices of the Royal Astronomical Society*, 20, 13, doi: [10.1093/mnras/20.1.13](https://doi.org/10.1093/mnras/20.1.13)

## Bibliography

- 1093/mnras/20.1.13 5
- Carroll, B. W., & Ostlie, D. A. 2007, *An Introduction to Modern Astrophysics*, 2nd edn. (San Francisco: Pearson Addison-Wesley) 3
- Carroll, T. A., & Staude, J. 2001, *Astronomy & Astrophysics*, 378, 316, doi: [10.1051/0004-6361:20011167](https://doi.org/10.1051/0004-6361:20011167) 171
- Cauzzi, G., Reardon, K. P., Uitenbroek, H., et al. 2008, *Astronomy & Astrophysics*, 480, 515, doi: [10.1051/0004-6361:20078642](https://doi.org/10.1051/0004-6361:20078642) 49
- Centeno, R., Schou, J., Hayashi, K., et al. 2014, *Solar Physics*, 289, 3531, doi: [10.1007/s11207-014-0497-7](https://doi.org/10.1007/s11207-014-0497-7) 56
- Cheung, M. C., Rempel, M., Chintzoglou, G., et al. 2019, *Nature Astronomy*, 3, 160, doi: [10.1038/s41550-018-0629-3](https://doi.org/10.1038/s41550-018-0629-3) 13, 135
- Clevert, D.-A., Unterthiner, T., & Hochreiter, S. 2015, 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings, 1. <https://arxiv.org/abs/1511.07289> 68
- Colella, P., & Woodward, P. R. 1984, *Journal of Computational Physics*, 54, 174, doi: [10.1016/0021-9991\(84\)90143-8](https://doi.org/10.1016/0021-9991(84)90143-8) 41
- Curtis, A. R., Powell, M. J., & Reid, J. K. 1974, *IMA Journal of Applied Mathematics (Institute of Mathematics and Its Applications)*, 13, 117, doi: [10.1093/imamat/13.1.117](https://doi.org/10.1093/imamat/13.1.117) 92
- Cybenko, G. 1989, *Mathematics of Control, Signals, and Systems*, 2, 303, doi: [10.1007/BF02551274](https://doi.org/10.1007/BF02551274) 65
- De Feiter, L. D., & Švestka, Z. 1975, *Solar Physics*, 41, 415, doi: [10.1007/BF00154078](https://doi.org/10.1007/BF00154078) 94, 107
- de la Cruz Rodríguez, J., Leenaarts, J., Danilovic, S., & Uitenbroek, H. 2019, *Astronomy & Astrophysics*, 623, A74. <https://arxiv.org/abs/1810.08441> 61, 193
- de la Cruz Rodríguez, J., Löfdahl, M. G., Sütterlin, P., Hillberg, T., & Rouppe van der Voort, L. 2015, *Astronomy & Astrophysics*, 573, A40, doi: [10.1051/0004-6361/201424319](https://doi.org/10.1051/0004-6361/201424319) 51
- de la Cruz Rodríguez, J., & Piskunov, N. 2013, *The Astrophysical Journal*, 764, 33, doi: [10.1088/0004-637X/764/1/33](https://doi.org/10.1088/0004-637X/764/1/33) 19, 76, 79
- de la Cruz Rodríguez, J., & van Noort, M. 2017, *Space Science Reviews*, 210, 109, doi: [10.1007/s11214-016-0294-8](https://doi.org/10.1007/s11214-016-0294-8) 49, 55, 60
- De Moortel, I., & Browning, P. 2015, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373, doi: [10.1098/rsta.2014.0269](https://doi.org/10.1098/rsta.2014.0269) 4
- del Toro Iniesta, J. C. 2003, *Introduction to Spectropolarimetry* (Cambridge, UK: Cambridge University Press) 58, 59

## Bibliography

- del Toro Iniesta, J. C., & Ruiz Cobo, B. 2016, *Living Reviews in Solar Physics*, 13, 1, doi: [10.1007/s41116-016-0005-2](https://doi.org/10.1007/s41116-016-0005-2) 55
- Díaz Baso, C. J., Asensio Ramos, A., & de la Cruz Rodríguez, J. 2021, arXiv e-prints, 2108.07089. <https://arxiv.org/abs/2108.07089> 62, 172
- Dinh, L., Krueger, D., & Bengio, Y. 2014, arXiv e-prints, arXiv:1410.8516. <https://arxiv.org/abs/1410.8516> 173
- Dinh, L., Sohl-Dickstein, J., & Bengio, S. 2016, arXiv e-prints, arXiv:1605.08803. <https://arxiv.org/abs/1605.08803> 173
- Dorfi, E., & Drury, L. 1987, *Journal of Computational Physics*, 69, 175, doi: [10.1016/0021-9991\(87\)90161-6](https://doi.org/10.1016/0021-9991(87)90161-6) 12, 86
- Dorfi, E. A. 1997, in *Computational Methods for Astrophysical Fluid Flow: Saas-Fee Advanced Course 27 (Berlin/Heidelberg: Springer-Verlag)*, 263–341, doi: [10.1007/3-540-31632-9\\_3](https://doi.org/10.1007/3-540-31632-9_3) 92
- Druett, M. K., & Zharkova, V. V. 2018, *Astronomy & Astrophysics*, 610, A68, doi: [10.1051/0004-6361/201731053](https://doi.org/10.1051/0004-6361/201731053) 12
- . 2019, *Astronomy & Astrophysics*, 623, A20, doi: [10.1051/0004-6361/201732427](https://doi.org/10.1051/0004-6361/201732427) 12
- Emslie, A. G. 1978, *The Astrophysical Journal*, 224, 241, doi: [10.1086/156371](https://doi.org/10.1086/156371) 14, 89
- Emslie, A. G., & Sturrock, P. A. 1982, *Solar Physics*, 80, 99, doi: [10.1007/BF00153426](https://doi.org/10.1007/BF00153426) 7
- Fabiani Bendicho, P., Trujillo Bueno, J., & Auer, L. 1997, *Astronomy & Astrophysics*, 324, 161 28, 35, 36
- Fang, C., Henoux, J., & Gan, W. 1993, *Astronomy & Astrophysics*, 274, 917 22, 89
- Feautrier, P. 1964, *SAO Special Report*, 167, 80 17
- Fisher, G., Canfield, R. C., & McClymont, A. N. 1985, *The Astrophysical Journal*, 289, 414 45
- Fletcher, L., & Hudson, H. S. 2008, *The Astrophysical Journal*, 675, 1645, doi: [10.1086/527044](https://doi.org/10.1086/527044) 7
- Fletcher, L., Dennis, B. R., Hudson, H. S., et al. 2011, *Space Science Reviews*, 159, 19, doi: [10.1007/s11214-010-9701-8](https://doi.org/10.1007/s11214-010-9701-8) 5, 48
- Fontenla, J., Avrett, E., & Loeser, R. 1993, *The Astrophysical Journal*, 406, 319 3, 78
- Fritsch, F. N., & Butland, J. 1984, *SIAM Journal on Scientific and Statistical Computing*, 5, 300, doi: [10.1137/0905021](https://doi.org/10.1137/0905021) 79
- Gafeira, R., Orozco Suárez, D., Milić, I., et al. 2021, *Astronomy & Astrophysics*, 651, A31, doi: [10.1051/0004-6361/201936910](https://doi.org/10.1051/0004-6361/201936910) 172
- Gayley, K. G. 1990, PhD thesis, California Univ., San Diego. 112

## Bibliography

- Geltman, S. 1962, *The Astrophysical Journal*, 136, 935, doi: [10.1086/147447](https://doi.org/10.1086/147447) 75
- Glorot, X., Bordes, A., & Bengio, Y. 2011, in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, ed. G. G. Dudík, D. Dunson, & Miroslav (Fort Lauderdale, FL, USA: PMLR), 315–323 68
- Godunov, S. 1959, *Matematicheskii Sbornik*, 47(89), 271 41
- Gouttebroze, P. 1986, *Astronomy & Astrophysics*, 160, 195 32
- Gouttebroze, P., & Heinzel, P. 2002, *Astronomy & Astrophysics*, 385, 273, doi: [10.1051/0004-6361:20020142](https://doi.org/10.1051/0004-6361:20020142) 101
- Graham, D. R., Hannah, I. G., Fletcher, L., & Milligan, R. O. 2013, *The Astrophysical Journal*, 767, 83, doi: [10.1088/0004-637X/767/1/83](https://doi.org/10.1088/0004-637X/767/1/83) 6
- Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., & Smola, A. 2012, *Journal of Machine Learning Research*, 13, 723 179
- Gropp, W., Lusk, E., Doss, N., & Skjellum, A. 1996, *Parallel Computing*, 22, 789, doi: [10.1016/0167-8191\(96\)00024-5](https://doi.org/10.1016/0167-8191(96)00024-5) 76
- Gudiksen, B. V., Carlsson, M., Hansteen, V. H., et al. 2011, *Astronomy & Astrophysics*, 531, A154, doi: [10.1051/0004-6361/201116520](https://doi.org/10.1051/0004-6361/201116520) 13, 120
- Gurman, J. B. 1992, in *NATO Advanced Study Institute "The Sun: a Laboratory for Astrophysics"* held in Crieff, Scotland on 16-29 June 1991, ed. J. T. Schmelz & J. C. Brown (Kluwer Academic), 245–260 3
- Gustafsson, B. 1973, *Uppsala Astronomical Observatory Annals*, 5, 1 94
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357, doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2) 9, 72
- He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv e-prints, arXiv:1512.03385. <https://arxiv.org/abs/1512.03385> 67
- Heinzel, P. 1995, *Astronomy & Astrophysics*, 299, 563 33
- Heinzel, P., Karlický, M., Kotrč, P., & Švestka, Z. 1994, *Solar Physics*, 152, 393, doi: [10.1007/BF00680446](https://doi.org/10.1007/BF00680446) 49
- Heinzel, P., Kašparová, J., Varady, M., Karlický, M., & Moravec, Z. 2015, *Proceedings of the International Astronomical Union*, 11, 233, doi: [10.1017/S1743921316000363](https://doi.org/10.1017/S1743921316000363) 12, 13
- Hirayama, T. 1974, *Solar Physics*, 34, 323, doi: [10.1007/BF00153671](https://doi.org/10.1007/BF00153671) 5
- Hong, J., Li, Y., Ding, M. D., & Carlsson, M. 2019, *The Astrophysical Journal*, 879, 128, doi: [10.3847/1538-4357/ab262e](https://doi.org/10.3847/1538-4357/ab262e) 93, 107
- Huang, W. F., Ren, Y. X., & Jiang, X. 2018, *Acta Mechanica Sinica/Lixue Xuebao*, 34, 37,

## Bibliography

- doi: [10.1007/s10409-017-0715-2](https://doi.org/10.1007/s10409-017-0715-2) 42
- Hubený, I. 1982, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 27, 593, doi: [10.1016/0022-4073\(82\)90052-8](https://doi.org/10.1016/0022-4073(82)90052-8) 31
- Hubený, I., & Mihalas, D. 2014, *Theory of Stellar Atmospheres* (Princeton, NJ: Princeton University Press) 3, 15, 26, 31, 75
- Hudson, H. S. 1972, *Solar Physics*, 24, 414, doi: [10.1007/BF00153384](https://doi.org/10.1007/BF00153384) 6
- Hudson, H. S., Strong, K. T., Dennis, B. R., et al. 1994, *The Astrophysical Journal*, 422, L25, doi: [10.1086/187203](https://doi.org/10.1086/187203) 6
- Hudson, H. S., Wolfson, C. J., & Metcalf, T. R. 2006, *Solar Physics*, 234, 79, doi: [10.1007/s11207-006-0056-y](https://doi.org/10.1007/s11207-006-0056-y) 135
- Hummer, D. G., & Rybicki, G. B. 1982, *The Astrophysical Journal*, 263, 925 82
- Hunter, J. D. 2007, *Computing in Science & Engineering*, 9, 90, doi: [10.1109/MCSE.2007.559](https://doi.org/10.1109/MCSE.2007.559)
- Ibgui, L., Hubeny, I., Lanz, T., & Stehlé, C. 2013, *Astronomy & Astrophysics*, 549, 1, doi: [10.1051/0004-6361/201220468](https://doi.org/10.1051/0004-6361/201220468) 129
- Ishizawa, T. 1971, *Publications of the Astronomical Society of Japan*, 23, 75 94
- Isobe, H., Kubo, M., Minoshima, T., et al. 2007, *Publications of the Astronomical Society of Japan*, 59, 807, doi: [10.1093/pasj/59.sp3.s807](https://doi.org/10.1093/pasj/59.sp3.s807) 135
- Ivanov, V. V., & Serbin, V. M. 1984, *Soviet Astronomy*, 28, 405 12
- Janett, G., Steiner, O., Alsina Ballester, E., Belluzzi, L., & Mishra, S. 2019, *Astronomy & Astrophysics*, 624, A104, doi: [10.1051/0004-6361/201834761](https://doi.org/10.1051/0004-6361/201834761) 41
- Janett, G., Steiner, O., & Belluzzi, L. 2018, *The Astrophysical Journal*, 865, 16, doi: [10.3847/1538-4357/aad4a4](https://doi.org/10.3847/1538-4357/aad4a4) 19, 79
- Jegou, S., Drozdal, M., Vazquez, D., Romero, A., & Bengio, Y. 2017, in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (IEEE)*, 1175–1183, doi: [10.1109/CVPRW.2017.156](https://doi.org/10.1109/CVPRW.2017.156) 68
- Jiang, G.-S., & Shu, C.-W. 1996, *Journal of Computational Physics*, 126, 202, doi: [10.1006/jcph.1996.0130](https://doi.org/10.1006/jcph.1996.0130) 42
- Jing, J., Xu, Y., Cao, W., et al. 2016, *Scientific Reports*, 6, 24319, doi: [10.1038/srep24319](https://doi.org/10.1038/srep24319) 135
- John, T. L. 1988, *Astronomy & Astrophysics*, 193, 189 75
- Johnson, L. C. 1972, *The Astrophysical Journal*, 174, 227, doi: [10.1086/151486](https://doi.org/10.1086/151486) 80
- Judge, P. G. 2005, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 92, 479, doi: [10.1016/j.jqsrt.2004.08.009](https://doi.org/10.1016/j.jqsrt.2004.08.009) 112

## Bibliography

- . 2017, *The Astrophysical Journal*, 851, 5, doi: [10.3847/1538-4357/aa96a9](https://doi.org/10.3847/1538-4357/aa96a9) 80, 81, 82
- Kašparová, J., Carlsson, M., Varady, M., & Heinzel, P. 2019, in *Radiative Signatures From the Cosmos*, ASP Conference Serries, Vol. 519, ed. K. Werner, C. Stehlé, T. Rauch, & T. M. Lanz (San Francisco: Astronomical Society of the Pacific), 141. <https://ui.adsabs.harvard.edu/abs/2019ASPC..519..141K> 13, 105
- Kašparová, J., Heinzel, P., Varady, M., & Karlický, M. 2003, in *Stellar Atmosphere Modeling*, ASP Conference Proceedings, Vol. 288, ed. I. Hubeny, D. Mihalas, & K. Werner (San Francisco: Astronomical Society of the Pacific), 544. <https://ui.adsabs.harvard.edu/abs/2003ASPC..288..544K> 34, 90
- Kerr, G. S., Allred, J. C., & Carlsson, M. 2019a, *The Astrophysical Journal*, 883, 57, doi: [10.3847/1538-4357/ab3c24](https://doi.org/10.3847/1538-4357/ab3c24) 32, 87, 96, 121
- Kerr, G. S., Allred, J. C., & Polito, V. 2020, *The Astrophysical Journal*, 900, 18, doi: [10.3847/1538-4357/abaa46](https://doi.org/10.3847/1538-4357/abaa46) 46, 87
- Kerr, G. S., Carlsson, M., & Allred, J. C. 2019b, *The Astrophysical Journal*, 885, 119, doi: [10.3847/1538-4357/ab48ea](https://doi.org/10.3847/1538-4357/ab48ea) 88, 96, 104, 121
- Kerr, G. S., Carlsson, M., Allred, J. C., Young, P. R., & Daw, A. N. 2019c, *The Astrophysical Journal*, 871, 23, doi: [10.3847/1538-4357/aaf46e](https://doi.org/10.3847/1538-4357/aaf46e) 88
- Kerr, G. S., Fletcher, L., Russell, A. J. B., & Allred, J. C. 2016, *The Astrophysical Journal*, 827, 101, doi: [10.3847/0004-637X/827/2/101](https://doi.org/10.3847/0004-637X/827/2/101) 15
- Kerr, G. S., Xu, Y., Allred, J. C., et al. 2021, *The Astrophysical Journal*, 912, 153, doi: [10.3847/1538-4357/abf42d](https://doi.org/10.3847/1538-4357/abf42d) 50
- Kingma, D., & Ba, J. 2014, arXiv e-prints. <https://arxiv.org/abs/1412.6980> 67, 178
- Kiselman, D., Pereira, T. M. D., Gustafsson, B., et al. 2011, *Astronomy & Astrophysics*, 535, A14, doi: [10.1051/0004-6361/201117553](https://doi.org/10.1051/0004-6361/201117553) 51
- Klimchuk, J. A. 2006, *Solar Physics*, 234, 41, doi: [10.1007/s11207-006-0055-z](https://doi.org/10.1007/s11207-006-0055-z) 4
- Klimchuk, J. A., Patsourakos, S., & Cargill, P. J. 2008, *The Astrophysical Journal*, 682, 1351, doi: [10.1086/589426](https://doi.org/10.1086/589426) 46
- Kopp, R., & Pneuman, G. 1976, *Solar Physics*, 50, 85, doi: [10.1007/BF00206193](https://doi.org/10.1007/BF00206193) 5
- Kowalski, A. F., Allred, J. C., Daw, A., Cauzzi, G., & Carlsson, M. 2017a, *The Astrophysical Journal*, 836, 12, doi: [10.3847/1538-4357/836/1/12](https://doi.org/10.3847/1538-4357/836/1/12) 121
- Kowalski, A. F., Hawley, S. L., Carlsson, M., et al. 2015, *Solar Physics*, 290, 3487, doi: [10.1007/s11207-015-0708-x](https://doi.org/10.1007/s11207-015-0708-x) 87
- Kowalski, A. F., Allred, J. C., Uitenbroek, H., et al. 2017b, *The Astrophysical Journal*, 837,

## Bibliography

- 125, doi: [10.3847/1538-4357/aa603e](https://doi.org/10.3847/1538-4357/aa603e) 15, 23, 62, 73
- Krucker, S., Battaglia, M., Cargill, P. J., et al. 2008, *The Astronomy and Astrophysics Review*, 16, 155, doi: [10.1007/s00159-008-0014-9](https://doi.org/10.1007/s00159-008-0014-9) 6
- Kunasz, P., & Auer, L. H. 1988, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 39, 67, doi: [10.1016/0022-4073\(88\)90021-0](https://doi.org/10.1016/0022-4073(88)90021-0) 123
- Kuridze, D., Henriques, V. M. J., Mathioudakis, M., et al. 2018, *The Astrophysical Journal*, 860, 10, doi: [10.3847/1538-4357/aac26d](https://doi.org/10.3847/1538-4357/aac26d) 49, 61
- Kuridze, D., Mathioudakis, M., Simões, P. J. A., et al. 2015, *The Astrophysical Journal*, 813, 125, doi: [10.1088/0004-637X/813/2/125](https://doi.org/10.1088/0004-637X/813/2/125) 15, 49, 50, 62, 164, 185, 192
- Kurucz, R. L., van Dishoeck, E. F., & Tarafdar, S. P. 1987, *The Astrophysical Journal*, 322, 992, doi: [10.1086/165794](https://doi.org/10.1086/165794) 75
- Kuzmanovska, O., Atanacković, O., & Faurobert, M. 2017, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 196, 230, doi: [10.1016/j.jqsrt.2017.04.009](https://doi.org/10.1016/j.jqsrt.2017.04.009) 28, 192
- Landi Degl'Innocenti, E., & Landi Degl'Innocenti, M. 1977, *Astronomy & Astrophysics*, 56, 111 59
- Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. 1998, *Proceedings of the IEEE*, 86, 2278 64
- Leenaarts, J., Carlsson, M., Hansteen, V., & Rutten, R. J. 2007, *Astronomy & Astrophysics*, 473, 625, doi: [10.1051/0004-6361:20078161](https://doi.org/10.1051/0004-6361:20078161) 13, 120
- Leenaarts, J., Carlsson, M., & Rouppe Van Der Voort, L. 2012a, *The Astrophysical Journal*, 749, doi: [10.1088/0004-637X/749/2/136](https://doi.org/10.1088/0004-637X/749/2/136) 50, 105, 107, 117, 120, 135
- Leenaarts, J., Pereira, T., & Uitenbroek, H. 2012b, *Astronomy & Astrophysics*, 543, A109, doi: [10.1051/0004-6361/201219394](https://doi.org/10.1051/0004-6361/201219394) 32
- Léger, L., Chevallier, L., & Paletou, F. 2007, *Astronomy & Astrophysics*, 470, 1, doi: [10.1051/0004-6361:20066730](https://doi.org/10.1051/0004-6361:20066730) 36
- LeVeque, R. J. 1997, in *Computational Methods for Astrophysical Fluid Flow: Saas-Fee Advanced Course 27* (Berlin/Heidelberg: Springer-Verlag), 1–159, doi: [10.1007/3-540-31632-9\\_1](https://doi.org/10.1007/3-540-31632-9_1) 36, 44
- . 2002, *Finite Volume Methods for Hyperbolic Problems* (Cambridge University Press), doi: [10.1017/CB09780511791253](https://doi.org/10.1017/CB09780511791253) 36
- Libbrecht, T., de la Cruz Rodríguez, J., Danilovic, S., Leenaarts, J., & Pazira, H. 2019, *Astronomy & Astrophysics*, 621, A35, doi: [10.1051/0004-6361/201833610](https://doi.org/10.1051/0004-6361/201833610) 50
- Liu, X.-D., Osher, S., & Chan, T. 1994, *Journal of Computational Physics*, 115, 200, doi: [10.1006/jcph.1994.1187](https://doi.org/10.1006/jcph.1994.1187) 42

## Bibliography

- Löfdahl, M. G., Hillberg, T., de la Cruz Rodríguez, J., et al. 2021, *Astronomy & Astrophysics*, 1, doi: [10.1051/0004-6361/202141326](https://doi.org/10.1051/0004-6361/202141326) 51
- Lu, Z., Pu, H., Wang, F., Hu, Z., & Wang, L. 2017, in *Advances in Neural Information Processing Systems 30*, ed. I. Guyon, U. V. Kuxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Long Beach, CA: Curran Associates, Inc.), 6231–6239. <https://arxiv.org/abs/1709.02540> 65, 66
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. 2013, in *ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, 28 68
- Mariska, J. T., Doschek, G. A., Boris, J. P., Oran, E. S., & Young, T. R., J. 1982, *The Astrophysical Journal*, 255, 783, doi: [10.1086/159877](https://doi.org/10.1086/159877) 12, 13
- Mariska, J. T., Emslie, A. G., & Li, P. 1989, *The Astrophysical Journal*, 341, 1067, doi: [10.1086/167564](https://doi.org/10.1086/167564) 13
- Martínez-Sykora, J., De Pontieu, B., Hansteen, V. H., et al. 2017, *Science*, 356, 1269, doi: [10.1126/science.aah5412](https://doi.org/10.1126/science.aah5412) 13
- Martínez-Sykora, J., Hansteen, V. H., Gudiksen, B., et al. 2019, *The Astrophysical Journal*, 878, 40, doi: [10.3847/1538-4357/ab1f0b](https://doi.org/10.3847/1538-4357/ab1f0b) 14
- McClymont, A. N., & Canfield, R. C. 1983, *The Astrophysical Journal*, 265, 483, doi: [10.1086/160692](https://doi.org/10.1086/160692) 12
- Mein, P. 1971, *Solar Physics*, 20, 3, doi: [10.1007/BF00146089](https://doi.org/10.1007/BF00146089) 59
- Mein, P., Mein, N., Malherbe, J.-M., et al. 1997, *Solar Physics*, 172, 161 49
- Metcalf, T. R., Canfield, R. C., Avrett, E. H., & Metcalf, F. T. 1990a, *The Astrophysical Journal*, 350, 463, doi: [10.1086/168400](https://doi.org/10.1086/168400) 59, 61, 113
- Metcalf, T. R., Canfield, R. C., & Saba, J. L. R. 1990b, *The Astrophysical Journal*, 365, 391, doi: [10.1086/169494](https://doi.org/10.1086/169494) 135
- Mihalas, D. 1978, *Stellar Atmospheres* (San Francisco: W.H. Freeman) 71, 75
- Milić, I., & Gafeira, R. 2020, *Astronomy & Astrophysics*, 644, A129, doi: [10.1051/0004-6361/201936537](https://doi.org/10.1051/0004-6361/201936537) 171, 172
- Milić, I., & van Noort, M. 2017, *Astronomy & Astrophysics*, 601, 1, doi: [10.1051/0004-6361/201629980](https://doi.org/10.1051/0004-6361/201629980) 61
- . 2018, *Astronomy & Astrophysics*, 617, A24, doi: [10.1051/0004-6361/201833382](https://doi.org/10.1051/0004-6361/201833382) 61, 77, 193
- Miller-Ricci, E., & Uitenbroek, H. 2002, *The Astrophysical Journal*, 566, 500, doi: [10.1086/337954](https://doi.org/10.1086/337954) 77
- Milligan, R. O., Kerr, G. S., Dennis, B. R., et al. 2014, *The Astrophysical Journal*, 793, 70,

## Bibliography

- doi: [10.1088/0004-637X/793/2/707](https://doi.org/10.1088/0004-637X/793/2/707)
- Moreno Cárdenas, F., Cristancho Sánchez, S., & Vargas Domínguez, S. 2016, *Advances in Space Research*, 57, 257, doi: [10.1016/j.asr.2015.08.026](https://doi.org/10.1016/j.asr.2015.08.026) 5
- Mrozek, T., & Tomczak, M. 2004, *Astronomy & Astrophysics*, 415, 377, doi: [10.1051/0004-6361:20034598](https://doi.org/10.1051/0004-6361:20034598) 6
- Muandet, K., Fukumizu, K., Sriperumbudur, B., & Schölkopf, B. 2017, *Foundations and Trends® in Machine Learning*, 10, 1, doi: [10.1561/2200000060](https://doi.org/10.1561/2200000060) 179
- Nagai, F. 1980, *Solar Physics*, 68, 351, doi: [10.1007/BF00156874](https://doi.org/10.1007/BF00156874) 12
- Nair, V., & Hinton, G. E. 2010, in *Proceedings of ICML*, Vol. 27, 807–814 65
- Neidig, D. F., Kiplinger, A. L., Cohl, H. S., & Wiborg, P. H. 1993, *The Astrophysical Journal*, 406, 306, doi: [10.1086/172442](https://doi.org/10.1086/172442) 135
- Nejezchleba, T. 1998, *Astronomy & Astrophysics*, 127, 607, doi: [10.1051/aas:1998118](https://doi.org/10.1051/aas:1998118) 90
- Ng, K. C. 1974, *The Journal of Chemical Physics*, 61, 2680, doi: [10.1063/1.1682399](https://doi.org/10.1063/1.1682399) 111
- Olson, G. L., Auer, L. H., & Buchler, J. R. 1986, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 35, 431, doi: [10.1016/0022-4073\(86\)90030-0](https://doi.org/10.1016/0022-4073(86)90030-0) 27
- Olson, G. L., & Kunasz, P. 1987, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 38, 325, doi: [10.1016/0022-4073\(87\)90027-6](https://doi.org/10.1016/0022-4073(87)90027-6) 18, 84
- Osborne, C. M. J. 2020, *Lightspinner*, v1.0.0, Zenodo, doi: [10.5281/zenodo.5228268](https://doi.org/10.5281/zenodo.5228268) 83
- . 2021a, *Lightweaver*, Zenodo, doi: [10.5281/zenodo.4066860](https://doi.org/10.5281/zenodo.4066860) 69, 189
- . 2021b, *MsLightweaver: Thesis Release*, v1.2.0, Zenodo, doi: [10.5281/zenodo.5484108](https://doi.org/10.5281/zenodo.5484108) 93
- . 2021c, *MsLightweaver2d: Thesis Release*, v1.0.0, Zenodo, doi: [10.5281/zenodo.5484184](https://doi.org/10.5281/zenodo.5484184) 140
- Osborne, C. M. J., & Armstrong, J. A. 2019, *Radynversion: Zenodo Release*, v1.0.2, Zenodo, doi: [10.5281/zenodo.5484254](https://doi.org/10.5281/zenodo.5484254) 172
- Osborne, C. M. J., Armstrong, J. A., & Fletcher, L. 2019, *The Astrophysical Journal*, 873, 128, doi: [10.3847/1538-4357/ab07b4](https://doi.org/10.3847/1538-4357/ab07b4) 62, 171
- Osborne, C. M. J., Heinzel, P., Kašparová, J., & Fletcher, L. 2021, *Monthly Notices of the Royal Astronomical Society*, 507, 1972, doi: [10.1093/mnras/stab2156](https://doi.org/10.1093/mnras/stab2156) 93
- Osborne, C. M. J., & Milić, I. 2021, *The Astrophysical Journal*, 917, 14, doi: [10.3847/1538-4357/ac02be](https://doi.org/10.3847/1538-4357/ac02be) 9, 15, 30, 34, 69, 73, 74, 189
- Paletou, F. 1995, *Astronomy & Astrophysics*, 302, 587 33
- Paletou, F., & Léger, L. 2007, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 103, 57, doi: [10.1016/j.jqsrt.2006.07.003](https://doi.org/10.1016/j.jqsrt.2006.07.003) 28

## Bibliography

- Paszke, A., Gross, S., Massa, F., et al. 2019, in *Advances in Neural Information Processing Systems* 32, ed. H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Curran Associates, Inc.), 8024–8035. <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf> 67, 69
- Pereira, T. M. D., & Uitenbroek, H. 2015, *Astronomy & Astrophysics*, 574, A3, doi: [10.1051/0004-6361/201424785](https://doi.org/10.1051/0004-6361/201424785) 77
- Polito, V., Testa, P., & De Pontieu, B. 2019, *The Astrophysical Journal*, 879, L17, doi: [10.3847/2041-8213/ab290b](https://doi.org/10.3847/2041-8213/ab290b) 87
- Poore, G. M. 2015, *Computational Science and Discovery*, 8, doi: [10.1088/1749-4699/8/1/014010](https://doi.org/10.1088/1749-4699/8/1/014010) 9
- Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, *The Astronomical Journal*, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f) 9
- Raschka, S. 2015, *Python Machine Learning* (Birmingham: Packt Publishing) 64, 65, 66
- Reep, J. W., Bradshaw, S. J., Crump, N. A., & Warren, H. P. 2019, *The Astrophysical Journal*, 871, 18, doi: [10.3847/1538-4357/aaf580](https://doi.org/10.3847/1538-4357/aaf580) 120
- Reep, J. W., & Russell, A. J. B. 2016, *The Astrophysical Journal*, 818, L20, doi: [10.3847/2041-8205/818/1/L20](https://doi.org/10.3847/2041-8205/818/1/L20) 15
- Rempel, M. 2016, *The Astrophysical Journal*, 834, 10, doi: [10.3847/1538-4357/834/1/10](https://doi.org/10.3847/1538-4357/834/1/10) 13, 45
- Rempel, M., Schüssler, M., & Knölker, M. 2009, *The Astrophysical Journal*, 691, 640, doi: [10.1088/0004-637X/691/1/640](https://doi.org/10.1088/0004-637X/691/1/640) 13
- Rimmele, T. R., Warner, M., Keil, S. L., et al. 2020, *Solar Physics*, 295, 172, doi: [10.1007/s11207-020-01736-7](https://doi.org/10.1007/s11207-020-01736-7) 52
- Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, *Astronomy & Astrophysics*, 558, A33, doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068) 9
- Rubio Da Costa, F., Fletcher, L., Labrosse, N., & Zuccarello, F. 2009, *Astronomy & Astrophysics*, 507, 1005, doi: [10.1051/0004-6361/200912651](https://doi.org/10.1051/0004-6361/200912651) 93
- Rubio da Costa, F., Kleint, L., Petrosian, V., Liu, W., & Allred, J. C. 2016, *The Astrophysical Journal*, 827, 38, doi: [10.3847/0004-637x/827/1/38](https://doi.org/10.3847/0004-637x/827/1/38) 15, 49, 62
- Ruiz Cobo, B., & del Toro Iniesta, J. 1992, *The Astrophysical Journal*, 398, 375 59, 60, 71
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986, *Learning representations by back-propagating errors*, Vol. 323 (Cambridge, MA, USA: MIT Press), 533–536, doi: [10.1038/323533a0](https://doi.org/10.1038/323533a0) 65, 66

## Bibliography

- Rusanov, V. V. 1962, *USSR Computational Mathematics and Mathematical Physics*, 1, 304, doi: [10.1016/0041-5553\(62\)90062-9](https://doi.org/10.1016/0041-5553(62)90062-9) 43
- Rybicki, G. 1972, in *Line Formation in the Presence of Magnetic Fields*, 145 27
- Rybicki, G. B. 1997, *The Astrophysical Journal*, 479, 357, doi: [10.1086/303877](https://doi.org/10.1086/303877) 21
- Rybicki, G. B., & Hummer, D. G. 1991, *Astronomy & Astrophysics*, 245, 171 13, 27, 85
- . 1992, *Astronomy & Astrophysics*, 262, 209 24, 27, 28, 77, 84
- Saint-Hilaire, P., Krucker, S., & Lin, R. P. 2008, *Solar Physics*, 250, 53, doi: [10.1007/s11207-008-9193-9](https://doi.org/10.1007/s11207-008-9193-9) 96
- Sánchez Almeida, J. 1992, *Solar Physics*, 137, 1, doi: [10.1007/BF00146572](https://doi.org/10.1007/BF00146572) 59
- Scharmer, G., & Carlsson, M. 1985, *Journal of Computational Physics*, 59, 56, doi: [10.1016/0021-9991\(85\)90107-X](https://doi.org/10.1016/0021-9991(85)90107-X) 12, 27
- Scharmer, G. B. 1981, *The Astrophysical Journal*, 249, 720, doi: [10.1086/159333](https://doi.org/10.1086/159333) 26
- Scharmer, G. B., Bjelksjo, K., Korhonen, T. K., Lindberg, B., & Petterson, B. 2003, in *Innovative Telescopes and Instrumentation for Solar Astrophysics*, ed. S. L. Keil & S. V. Avakyan, Vol. 4853, 341, doi: [10.1117/12.460377](https://doi.org/10.1117/12.460377) 51
- Scharmer, G. B., Löfdahl, M. G., Sliopen, G., & de la Cruz Rodríguez, J. 2019, *Astronomy & Astrophysics*, 626, A55, doi: [10.1051/0004-6361/201935735](https://doi.org/10.1051/0004-6361/201935735) 51
- Scharmer, G. B., Narayan, G., Hillberg, T., et al. 2008, *The Astrophysical Journal*, 689, L69, doi: [10.1086/595744](https://doi.org/10.1086/595744) 51
- Segura, A., Walkowicz, L. M., Meadows, V., Kasting, J., & Hawley, S. 2010, *Astrobiology*, 10, 751, doi: [10.1089/ast.2009.0376](https://doi.org/10.1089/ast.2009.0376) 1
- Shu, C.-W., & Osher, S. 1988, *Journal of Computational Physics*, 77, 439, doi: [10.1016/0021-9991\(88\)90177-5](https://doi.org/10.1016/0021-9991(88)90177-5) 44
- Simard, P., Steinkraus, D., & Platt, J. 2003, in *ICDAR '03, Vol. 1, Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings.* (Washington, DC, USA: IEEE Comput. Soc), 958–963, doi: [10.1109/ICDAR.2003.1227801](https://doi.org/10.1109/ICDAR.2003.1227801) 65
- Simões, P. J. A., Graham, D. R., & Fletcher, L. 2015, *Solar Physics*, 290, 3573, doi: [10.1007/s11207-015-0709-9](https://doi.org/10.1007/s11207-015-0709-9) 6
- Simões, P. J. A., Kerr, G. S., Fletcher, L., et al. 2017, *Astronomy & Astrophysics*, 605, A125, doi: [10.1051/0004-6361/201730856](https://doi.org/10.1051/0004-6361/201730856) 15, 62
- Simões, P. J. A., & Kontar, E. P. 2013, *Astronomy & Astrophysics*, 551, A135, doi: [10.1051/0004-6361/201220304](https://doi.org/10.1051/0004-6361/201220304) 7
- Socas-Navarro, H., De La Cruz Rodríguez, J., Asensio Ramos, A., Trujillo Bueno, J., & Ruiz

## Bibliography

- Cobo, B. 2015, *Astronomy & Astrophysics*, 577, 1, doi: [10.1051/0004-6361/201424860](https://doi.org/10.1051/0004-6361/201424860) 61, 71
- Socas-Navarro, H. 2005, *The Astrophysical Journal*, 621, 545, doi: [10.1086/427431](https://doi.org/10.1086/427431) 171
- Socas-Navarro, H., & Trujillo Bueno, J. 1997, *The Astrophysical Journal*, 490, 383, doi: [10.1086/304873](https://doi.org/10.1086/304873) 27, 85, 86
- Socas-Navarro, H., Trujillo Bueno, J., & Ruiz Cobo, B. 2000, *The Astrophysical Journal*, 530, 977, doi: [10.1086/308414](https://doi.org/10.1086/308414) 61
- Solanki, S. K. 2004, *Proceedings of the International Astronomical Union*, 2004, 195, doi: [10.1017/S1743921304005587](https://doi.org/10.1017/S1743921304005587) 3
- Sollum, E. 1999, PhD thesis, University of Oslo 12, 13, 120
- Spitzer, L., & Härm, R. 1953, *Physical Review*, 89, 977, doi: [10.1103/PhysRev.89.977](https://doi.org/10.1103/PhysRev.89.977) 45, 135
- Sriperumbudur, B. K., Fukumizu, K., Gretton, A., Lanckriet, G. R. G., & Schölkopf, B. 2009, in *Advances in Neural Information Processing Systems 22*, ed. Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Vancouver: Curran Associates, Inc.), 1750–1758. <https://proceedings.neurips.cc/paper/2009/file/685ac8cad1be5ac98da9556bc1c8d9e-Paper.pdf> 179, 180
- Steffen, M. 1990, *Astronomy & Astrophysics*, 239, 443 79
- Steiner, O., Züger, F., & Belluzzi, L. 2016, *Astronomy & Astrophysics*, 586, A42, doi: [10.1051/0004-6361/201527158](https://doi.org/10.1051/0004-6361/201527158) 19
- Štěpán, J., Bestard, J. J., & Bueno, J. T. 2020, *Astronomy & Astrophysics*, 636, 1, doi: [10.1051/0004-6361/202037566](https://doi.org/10.1051/0004-6361/202037566) 124, 131, 137
- Štěpán, J., & Trujillo Bueno, J. 2013, *Astronomy & Astrophysics*, 557, A143, doi: [10.1051/0004-6361/201321742](https://doi.org/10.1051/0004-6361/201321742) 19, 125, 126, 130
- Stilley, J. L., & Callaway, J. 1970, *The Astrophysical Journal*, 160, 245, doi: [10.1086/150423](https://doi.org/10.1086/150423) 75
- Strang, G. 1968, *SIAM Journal on Numerical Analysis*, 5, 506, doi: [10.1137/0705041](https://doi.org/10.1137/0705041) 44
- Sturrock, P. A. 1966, *Nature*, 211, 695, doi: [10.1038/211695a0](https://doi.org/10.1038/211695a0) 5
- Sui, L., Holman, G. D., & Dennis, B. R. 2007, *The Astrophysical Journal*, 670, 862, doi: [10.1086/522198](https://doi.org/10.1086/522198) 96
- Sutton, K. 1978, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 20, 333, doi: [10.1016/0022-4073\(78\)90102-4](https://doi.org/10.1016/0022-4073(78)90102-4) 23
- Švestka, Z. 1966, *Space Science Reviews*, 5, 388, doi: [10.1007/BF02653250](https://doi.org/10.1007/BF02653250) 49
- Syrovatskii, S. I., & Shmeleva, O. P. 1972, *Soviet Astronomy*, 16, 273 12

## Bibliography

- Tandberg-Hanssen, E., & Emslie, A. G. 1988, *The Physics of Solar Flares* (Cambridge, UK: Cambridge University Press) 7
- Tarafdar, S. P., & Vardya, M. S. 1973, *Monthly Notices of the Royal Astronomical Society*, 163, 261, doi: [10.1093/mnras/163.3.261](https://doi.org/10.1093/mnras/163.3.261) 75
- Tolstikhin, I., Bousquet, O., Gelly, S., & Schoelkopf, B. 2017, arXiv e-prints, arXiv:1711.01558. <https://arxiv.org/abs/1711.01558> 181
- Trujillo Bueno, J., & Fabiani Bendicho, P. 1995, *The Astrophysical Journal*, 455, 646, doi: [10.1086/176612](https://doi.org/10.1086/176612) 28
- Uitenbroek, H. 1989, *Astronomy & Astrophysics*, 216, 310 32
- . 2001, *The Astrophysical Journal*, 557, 389, doi: [10.1086/321659](https://doi.org/10.1086/321659) 24, 27, 28, 31, 32, 36, 71, 77, 84, 94, 108, 110
- . 2002, *The Astrophysical Journal*, 565, 1312, doi: [10.1086/324698](https://doi.org/10.1086/324698) 110, 121, 122
- Uitenbroek, H. 2006, in *Solar MHD Theory and Observations: A High Spatial Resolution Perspective*, ASP Conference Series, Vol. 354, ed. H. Uitenbroek, J. W. Leibacher, & R. F. Stein (San Francisco: Astronomical Society of the Pacific), 313 60
- van Driel-Gesztelyi, L., & Green, L. M. 2015, *Living Reviews in Solar Physics*, 12, 1, doi: [10.1007/lrsp-2015-1](https://doi.org/10.1007/lrsp-2015-1) 5
- van Leer, B. 1979, *Journal of Computational Physics*, 32, 101, doi: [10.1016/0021-9991\(79\)90145-1](https://doi.org/10.1016/0021-9991(79)90145-1) 41, 92
- van Noort, M., Der Voort, L. R. V., & Löfdahl, M. G. 2005, *Solar Physics*, 228, 191, doi: [10.1007/s11207-005-5782-z](https://doi.org/10.1007/s11207-005-5782-z) 51
- van Noort, M., Hubeny, I., & Lanz, T. 2002, *The Astrophysical Journal*, 568, 1066, doi: [10.1086/338949](https://doi.org/10.1086/338949) 129
- Varady, M., Kašparová, J., Moravec, Z., Heinzl, P., & Karlický, M. 2010, *IEEE Transactions on Plasma Science*, 38, 2249, doi: [10.1109/TPS.2010.2057449](https://doi.org/10.1109/TPS.2010.2057449) 12, 13, 14
- Vernazza, J. E., Avrett, E. H., & Loeser, R. 1981, *The Astrophysical Journal Supplement Series*, 45, 635, doi: [10.1086/190731](https://doi.org/10.1086/190731) 3, 96, 105, 176
- Viallet, M., Baraffe, I., & Walder, R. 2011, *Astronomy & Astrophysics*, 531, A86, doi: [10.1051/0004-6361/201016374](https://doi.org/10.1051/0004-6361/201016374) 43
- Victor, G. A., & Dalgarno, A. 1969, *The Journal of Chemical Physics*, 50, 2535, doi: [10.1063/1.1671412](https://doi.org/10.1063/1.1671412) 75
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2) 9

## Bibliography

- Vissers, G. J. M., Danilovic, S., de la Cruz Rodríguez, J., et al. 2021, *Astronomy & Astrophysics*, 645, A1, doi: [10.1051/0004-6361/202038900](https://doi.org/10.1051/0004-6361/202038900) 50
- Wang, H., Gary, D. E., Zirin, H., et al. 1995, *The Astrophysical Journal*, 453, 505–49
- Waskom, M. 2021, *Journal of Open Source Software*, 6, 3021, doi: [10.21105/joss.03021](https://doi.org/10.21105/joss.03021) 9
- Wedemeyer-Böhm, S., & Carlsson, M. 2011, *Astronomy & Astrophysics*, 528, A1, doi: [10.1051/0004-6361/201016186](https://doi.org/10.1051/0004-6361/201016186) 105
- Woods, T. N., Kopp, G., & Chamberlin, P. C. 2006, *Journal of Geophysical Research*, 111, A10S14, doi: [10.1029/2005JA011507](https://doi.org/10.1029/2005JA011507) 2
- Xu, Y., Cao, W., Liu, C., et al. 2006, *The Astrophysical Journal*, 641, 1210, doi: [10.1086/500632](https://doi.org/10.1086/500632) 135
- Yadav, R., Díaz Baso, C. J., de la Cruz Rodríguez, J., Calvo, F., & Morosin, R. 2021, *Astronomy & Astrophysics*, 649, A106, doi: [10.1051/0004-6361/202039857](https://doi.org/10.1051/0004-6361/202039857) 50
- Zeng, Z., Qiu, J., Cao, W., & Judge, P. G. 2014, *The Astrophysical Journal*, 793, 87, doi: [10.1088/0004-637X/793/2/87](https://doi.org/10.1088/0004-637X/793/2/87) 50
- Zhu, Y., Kowalski, A. F., Tian, H., et al. 2019, *The Astrophysical Journal*, 879, 19, doi: [10.3847/1538-4357/ab2238](https://doi.org/10.3847/1538-4357/ab2238) 121
- Zirin, H. 1992, in *NATO Advanced Study Institute "The Sun: a Laboratory for Astrophysics"* held in Crieff, Scotland on 16-29 June 1991, ed. J. T. Schmelz & J. C. Brown (Kluwer Academic), 175–190 2