Where do the heavy metals of the Universe come from? The answer lies in outer space, when two neutron stars collide. This violent event creates electromagnetic emission across the spectrum, powered by the decay of unstable heavy elements, which we call a kilonova. In this thesis, the modelling of kilonovae by radiative transfer is explored, focussing on the late-time emission in the nebular phase. The first part of this work introduces the field both from a physical and technical perspective. The papers conducted for this thesis are then attached, ending with the presentation of some of the first fully self-consistent KN spectra generated in the nebular phase.
Non-Local Thermodynamic Equilibrium Spectral Modelling of Kilonovae

Quentin Pognan

Academic dissertation for the Degree of Doctor of Philosophy in Astronomy at Stockholm University to be publicly defended on Monday 18 December 2023 at 10.00 in sal FB52, AlbaNova universitetscentrum, Roslagstullsbacken 21 and online via Zoom, public link is available at the department website.

Abstract

The astrophysical origin of rapid neutron capture (r-process) elements has long remained a puzzle and been the object of scientific debate. Neutron star (NS) mergers have historically been suggested as an ideal site for the creation of these elements, and were propelled into focus following the detection of the first binary neutron star (BNS) merger in 2017. The gravitational wave (GW) signal GW170817 was accompanied by a short gamma-ray burst (sGRB) GRB170817A, and a radioactively powered electromagnetic (EM) transient AT2017gfo, known as a kilonova (KN). Since this detection, the study of NS mergers has greatly expanded across the diverse fields that model the various stages of the merger, from GW signal modelling, to radiative transfer studies predicting the emergent KN lightcurves (LCs) and spectra.

One main goal of studying NS mergers and the associated KNe is to establish the importance of compact object mergers as key sites of r-process nucleosynthesis in the Universe. As such, identification of elements and their abundances within the merger ejecta represents a critical objective. LC and spectral analyses of KNe provide promising channels to do so, and require detailed models in order to interpret observational data. With complete GW and multi-band EM data only available for a single object thus far, the importance of detailed models regarding every aspect of KN physics is paramount. KN simulations typically make use of radiative transfer (RT) codes that propagate photons through the expanding ejecta, in order to provide LC and spectral outputs. These often model the early, photospheric times of the KN, when the ejecta are still dense enough such that the gas state is well described by Local Thermodynamic Equilibrium (LTE) conditions, which requires thermal collisional processes to dominate within the ejecta.

Since the ejecta are expanding rapidly however, these conditions cease to apply after several days, and the KN transitions to the Non-Local Thermodynamic Equilibrium (NLTE) regime, where thermal collisional processes are no longer dominant in establishing the gas state of the ejecta. This now requires the detailed modelling of various NLTE processes which increases the complexity, yet modelling of this regime can also provide great rewards. Notably, as times goes on and the ejecta continue to expand, they will eventually become optically thin to most wavelengths and enter the nebular phase. There, the spectra are expected to be emission line dominated, providing an excellent opportunity for element identification by spectral analysis.

This doctoral thesis conducts RT modelling in order to explore the NLTE regime of the KN in a systematic, physically accurate way. To this end, the spectral synthesis code SUMO (SUpernova MOnte Carlo Code) was adapted to model KNe, and used to investigate the spectral emission in the NLTE regime. The work in this doctoral thesis provides a first step into fully consistent modelling and analysis of KNe at later times, and a solid foundation from which to move forwards.

Keywords: Transients: Neutron Star Mergers, Radiative Transfer.

Stockholm 2023
http://urn.kb.se/resolve?urn=urn:nbn:se:diva-223538

ISBN 978-91-8014-583-1
ISBN 978-91-8014-584-8

Department of Astronomy
Stockholm University, 106 91 Stockholm
NON-LOCAL THERMODYNAMIC EQUILIBRIUM SPECTRAL MODELLING OF KILONOVAE

Quentin Pognan
Non-Local Thermodynamic Equilibrium Spectral Modelling of Kilonovae

Quentin Pognan
The astrophysical origin of rapid neutron capture (r-process) elements has long remained a puzzle and been the object of scientific debate. Neutron star (NS) mergers have historically been suggested as an ideal site for the creation of these elements, and were propelled into focus following the detection of the first binary neutron star (BNS) merger in 2017. The gravitational wave (GW) signal GW170817 was accompanied by a short gamma-ray burst (sGRB) GRB170817A, and a radioactively powered electromagnetic (EM) transient AT2017gfo, known as a kilonova (KN). Since this detection, the study of NS mergers has greatly expanded across the diverse fields that model the various stages of the merger, from GW signal modelling, to radiative transfer studies predicting the emergent KN lightcurves (LCs) and spectra.

One main goal of studying NS mergers and the associated KNe is to establish the importance of compact object mergers as key sites of r-process nucleosynthesis in the Universe. As such, identification of elements and their abundances within the merger ejecta represents a critical objective. LC and spectral analyses of KNe provide promising channels to do so, and require detailed models in order to interpret observational data. With complete GW and multi-band EM data only available for a single object thus far, the importance of detailed models regarding every aspect of KN physics is paramount. KN simulations typically make use of radiative transfer (RT) codes that propagate photons through the expanding ejecta in order to provide LC and spectral outputs. These often model the early, photospheric times of the KN, when the ejecta are still dense enough such that the gas state is well described by Local Thermodynamic Equilibrium (LTE) conditions, which require thermal collisional processes to dominate within the ejecta.

Since the ejecta are expanding rapidly however, these conditions cease to apply after several days, and the KN transitions to the Non-Local Thermodynamic Equilibrium (NLTE) regime, where thermal collisional processes are no longer dominant in establishing
the gas state of the ejecta. This now requires the detailed modelling of various NLTE processes which increases the complexity, yet modelling of this regime can also provide great rewards. Notably, as times goes on and the ejecta continue to expand, they will eventually become optically thin to most wavelengths and enter the nebular phase. There, the spectra are expected to be emission line dominated, providing an excellent opportunity for element identification by spectral analysis.

This PhD thesis conducts NLTE RT modelling in order to explore the NLTE regime of the KN in a systematic, physically accurate way. To this end, the spectral synthesis code SUMO (SUpernova MOnte Carlo Code) was adapted to model KNe, and used to investigate the spectral emission in the NLTE regime. As part of the work in this PhD thesis, the applicability of two assumptions commonly applied to the gas state of the ejecta, notably that of using the LTE condition, and the steady-state condition in the NLTE regime, were initially examined. The range of validity and consequences of incorrect application of these simplifications were quantitatively established. The subsequent work of this thesis provided the first full NLTE spectra of KNe conducted without the assumption of an optically thin medium. This work identified which elements produce signatures in late-time KN spectra, and provides a first step into fully consistent modelling and analysis of KNe in the NLTE regime.
Abstrakt - Svenska

Det har sedan lång tid funnits osäkerhet kring det astrofysikaliska ursprunget av de så kallade “r-process” grundämnen. Sammanslagningar (mergers) av neutronstjärnor (NS) har setts som en lovande källa, en teori som fick mer uppmärksamhet när den första NS-mergern observerades år 2017. En gravitationsvåg (GW) signal, GW170817, följdes av ett kort gamma-strålningsutbrott (sGRB), GRB170817A, och av en radioaktivt driven elektromagnetisk (EM) transient, AT2017gfo, också känd som en kilonova (KN). Efter den här upptäckten har forskningen kring NS-mergers blivit mer och mer omfattande, och täcker idag modellering från och med GW-signalen hela vägen till kilonovans sena spektra.

Ett av huvudmålen med forskningen av NS mergers och dess kilonovor är att bestämma hur viktiga de är som r-process källor, sett mot hela universums ämnesberikning över tid. På grund av det är det essentiellt att bestämma vilka ämnen finns i mergerns ejekta, och i vilka mängder. Studier av kilonovors ljuskurvor och spektra ger bra möjligheter till sådan analys, under förutsättning att det finns detaljerade modeller som kan jämföras med den observationella datan. I och med att det bara finns komplett GW och multi-band EM data för ett enda objekt än så länge, är behovet av detaljerade modeller för varje komponent av fysiken bakom kilonovor ännu större. KN simuleringar använder strålningstransport (RT) koder som sprider fotonerna genom det expanderande ejektat för att få fram observabler. De koderna modellerar ofta de tidiga, fotosfäriska epokerna av kilonovan, när ejektat fortfarande är tillräckligt tätt för att beskrivas med ett antagande om Local Thermodynamic Equilibrium (LTE), vilket kräver att termiska kollisionssprocesser dominerar.

Eftersom en kilonova expanderar snabbt, så är inte det här antagandet tillfredsställande efter några dagar, utan kilonovan kräver då istället en Non-Local Thermodynamic Equilibrium (NLTE) behandling. Trots att den här modelleringen är mer komplex än i LTE, så kan den ge oss ny och viktig information. Till exempel blir kilonovan under den nebulära
Contents

I Abstract - English i
II Abstract - Swedish iii
III List of Papers ix
IV Author’s Contribution xi
V List of Figures xv
VI Acronyms xvii

1 Introduction 1

2 Progenitor Systems and Gravitational Waves 7
  2.1 Compact Object Remnants . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
  2.2 Binary Formation and Evolution . . . . . . . . . . . . . . . . . . . . . . . . 8
    2.2.1 Common Envelope Evolution . . . . . . . . . . . . . . . . . . . . . 10
    2.2.2 Chemically Homogeneous Evolution . . . . . . . . . . . . . . . . . 11
    2.2.3 Dynamical Interactions . . . . . . . . . . . . . . . . . . . . . . . . . 12
  2.3 Gravitational Wave Signals . . . . . . . . . . . . . . . . . . . . . . . . . . 13

3 Ejecta and r-process Nucleosynthesis 21
  3.1 Basics of the Rapid Neutron Capture Process . . . . . . . . . . . . . . . 21
  3.2 Properties of Compact Binary Merger Ejecta . . . . . . . . . . . . . . . . 26
    3.2.1 Progenitor Properties and Merger Remnants . . . . . . . . . . . . . 28
    3.2.2 Dynamical Ejecta . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
3.2.3 Disc Ejecta ........................................... 31
3.3 Energy Deposition ...................................... 32
  3.3.1 $\beta$-decay ...................................... 34
  3.3.2 $\alpha$-decay and Fission .......................... 37

4 Electromagnetic Emission 39
  4.1 Gamma Ray Bursts .................................. 40
    4.1.1 GRB Afterglows ................................ 43
  4.2 The Kilonova ........................................ 45
    4.2.1 Opacity ......................................... 46
    4.2.2 Lightcurves and SEDs ............................ 49
    4.2.3 Spectral Analyses of AT2017gfo ................. 54

5 The Nebular Phase 63
  5.1 Physics of Nebular Phase Kilonovae .................. 64
    5.1.1 Temperature ...................................... 65
    5.1.2 Ionisation ........................................ 73
    5.1.3 Excitation ........................................ 81
    5.1.4 The Time-Dependent Nebular Phase ................ 84

6 Kilonova Simulations with SUMO 85
  6.1 Gas State Solutions .................................. 85
  6.2 Radiative Transfer ................................... 87
  6.3 Atomic Data for Kilonova Modelling .................. 90
  6.4 Numerical Tests and Code Limitations ............... 92
    6.4.1 NLTE Solutions: The Newton-Raphson Method .... 92
    6.4.2 Atomic Data Limitations ......................... 95
    6.4.3 Resolutions and Ranges .......................... 98

7 Kilonova Modelling Techniques 105
  7.1 TARDIS .............................................. 106
  7.2 ARTIS ............................................... 109
  7.3 SEDONA .............................................. 112
  7.4 TH13 .................................................. 114
7.5 NLTE Studies .............................................................. 116

8 Paper Summary ......................................................... 121
8.1 Paper I: On the Validity of Steady-State for Nebular Phase Kilonovae . . . 121
8.2 Paper II: NLTE Effects on Kilonova Expansion Opacities .................. 122
8.3 Paper III: NLTE Spectra of Kilonovae ................................ 124

9 Summary and Future Directions .................................... 129

Acknowledgements ...................................................... 131

References
Papers in the Doctoral Thesis:

Paper I:
“On the Validity of Steady-State for Nebular Phase Kilonovae”,

Paper II:
“NLTE Effects on Kilonova Expansion Opacities”,

Paper III:
"NLTE Spectra of Kilonovae",

A summary of the papers and their place in the context of the research field can be found in Chapter 8. The papers themselves can be found at the end of this thesis, in their chronological order of publication. The reprints are authorised under the terms of the Creative Commons Attribution License, (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Contributions to the Papers

Author’s Contribution to Paper I:

- Chose the elemental composition of the model, and corresponding abundances.

- Implemented a new formula for energy deposition relevant to kilonovae in the SUMO code.

- Added the r-process elements Te, Ce, Pt, and Th into SUMO. Atomic data calculations by Jon Grumer from Uppsala University.

- Helped supervisor derive the analytical estimates for key timescales relevant to the study.

- Ran the models across the entire parameter space.

- Ran all additional tests related to the models, e.g. time-step testing for numerical purposes.

- Analysed the results with supervisor Anders Jerkstrand, and made all the plots.

- Wrote most of the paper. Appendices A and F, as well as parts of main text written and edited by Anders Jerkstrand. Atomic data section and Appendix B partially written by Jon Grumer.

Author’s Contribution to Paper II:

- Made Saha-Boltzmann solver for LTE models.

- Ran all the models relevant to the study.

- Analysed results alongside supervisor Anders Jerkstrand, and made all the plots.

- Wrote most of paper, some parts written and edited by Anders Jerkstrand.

Author’s Contribution to Paper III:

- Designed the model composition and calculated the energy deposition from nuclear network files provided by Shinya Wanajo.
• Conducted various numerical tests to decide on zone number, velocity spacing, density profile, spectral resolution etc. with guidance of supervisor.

• Implemented new complete r-process atomic data set from Jon Grumer into SUMO.

• Calibrated key atomic species Sr II and Y I. Y I was calibrated after first results were analysed.

• Ran all the models, including reruns after calibrating Y I, and correcting energy deposition issue for low $Y_e$ model.

• Analysed results alongside Anders Jerkstrand and Jon Grumer, and made corresponding plots.

New and Reused Material from Licentiate Thesis

I briefly describe here the material that has been reused from the Licentiate Thesis, as well as that which has been updated or added. Every figure has the origin written at the end of the caption, whether taken from a publication, reused from the Licentiate Thesis, or newly created for this doctoral thesis, marked as 'own figure'.

- **Abstract**: The abstract has been entirely rewritten to be consistent with the requirements for the PhD thesis, and to include text relevant to the third paper published since the Licentiate thesis.

- **Chapter 1**: The introduction has been reworked, with updates to include more recent works in the field. A paragraph has been added briefly describing the third paper included in the PhD Thesis. The description of what is included in this thesis has been updated to include the new chapters.

- **Chapter 2**: This chapter on progenitor systems and gravitational waves has been reused, with some updates to include the latest results of gravitational wave observatory results from the third observing run O3. Rate estimates for neutron star mergers have also been updated. Two additional binary formation channels have been briefly described.

- **Chapter 3**: Chapter 3 on the ejecta properties and r-process nucleosynthesis has also been reused. References to more modern works, as well as updates of binary neutron star merger rate estimates have been added.

- **Chapter 4**: Chapter 4 on the electromagnetic emission associated to merging neutron stars has been updated to include the latest observations of long gamma-ray burst afterglows emission excesses, which have been linked to r-process decay and thus provide signatures of kilonovae. In particular, the observation of GRB230723A is addressed, as this represents (potentially) the first observation of a kilonova with the James Webb Space Telescope, and the first spectra in the nebular phase. The spectral properties of AT2017gfo have also been added in more detail, in particular with respect to studies that have used photospheric spectra to identify elements in the ejecta. This is now also particularly relevant to Paper III, where we suggest a
transition from Rb I as a potential candidate for a P-Cygni feature seen around 5 days in the spectra of AT2017gfo.

- **Chapter 5**: This chapter on the physics of the nebular phase has been kept almost entirely as in the licentiate thesis. Since this chapter describes how the thermodynamic conditions are solved for physically, i.e. not considering numerical schemes, this was fully covered in the Licentiate Thesis which pertained to Papers I and II. A paragraph on thermal collisional ionisation, which was included in Paper III, has also been added.

- **Chapter 6**: Chapter 6 has been entirely rewritten. It now details the numerical functioning of SUMO. First, the thermodynamics are considered, i.e. the solving of the equations described in Chapter 5 from an operational perspective. The radiative transfer component is then addressed in broad terms detailing how the Monte Carlo method is implemented in SUMO. Since this was not changed from the original functioning, the focus is on the particular aspects which make SUMO stand out from other similar codes in the transient radiative transfer field. This chapter ends with a detailed section on the various code limitations and numerical tests carried out in the context of this thesis and kilonova modelling.

- **Chapter 7**: This chapter is entirely new, and describes the functioning of other codes used to model kilonovae. Their functioning is examined with key similarities and differences to SUMO highlighted. Four distinct codes are considered, and an additional unnamed code used for NLTE modelling in Hotokezaka et al. (2021) is also examined. The chapter ends with a paragraph summarising the general difficulties of accurate modelling of kilonovae in the NLTE regime, providing a rationale for why so few NLTE studies currently exist.

- **Chapter 8**: The paper summary chapter has been extended to include the summary of Paper III. The summary of Papers I and II is unchanged.

- **Chapter 9**: The outlook chapter summarising the thesis work and considering future directions has been rewritten in the current context of the thesis research and present state of the field.
List of Figures

2.1 Evolved Stellar Structure ........................................ 8
2.2 LIGO Interferometer .............................................. 14
2.3 GW170817 ...................................................... 15
2.4 LIGO Merger Masses .............................................. 17
2.5 GW170817 Sky Location ......................................... 18

3.1 Solar r-process Abundances ..................................... 23
3.2 Y_e Effect on r-process Yields .................................. 24
3.3 KN Ejecta Schematic ............................................ 27
3.4 Decay Product Thermalisation .................................. 34

4.1 GRB 170817A Lightcurve ....................................... 41
4.2 GRB 170817A Afterglow Lightcurves .......................... 44
4.3 GRB230723A JWST Spectra .................................... 45
4.4 KN AT2017gfo Lightcurves ..................................... 49
4.5 Blackbody Temperature fits of AT2017gfo ..................... 51
4.6 First Spectral Analysis of AT2017gfo ......................... 55
4.7 Two Component Spectral Fit to AT2017gfo .................... 56
4.8 Sr II in AT2017gfo ............................................. 57
4.9 Te III in AT2017gfo ............................................ 60

5.1 Approximation for Temperature Time Derivative in Steady-State .... 68
5.2 Heating and Cooling Timescales ................................ 72
5.3 Approximation for Ionisation Time Derivative in Steady-State ........ 75
5.4 Ionisation and Recombination Timescales ....................... 80
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>SUMO Schematic</td>
<td>86</td>
</tr>
<tr>
<td>6.2</td>
<td>Convergence Test</td>
<td>94</td>
</tr>
<tr>
<td>6.3</td>
<td>Branching Ratio Cut Test</td>
<td>97</td>
</tr>
<tr>
<td>6.4</td>
<td>Effect of $\tau_{\text{thresh}}$ on Emergent Spectra</td>
<td>99</td>
</tr>
<tr>
<td>6.5</td>
<td>Spatial Resolution Test</td>
<td>100</td>
</tr>
<tr>
<td>6.6</td>
<td>Spectral Resolution Test</td>
<td>101</td>
</tr>
<tr>
<td>6.7</td>
<td>Wavelength Range for Radiative Transfer Test</td>
<td>102</td>
</tr>
<tr>
<td>7.1</td>
<td>TARDIS Schematic</td>
<td>107</td>
</tr>
<tr>
<td>7.2</td>
<td>TARDIS Spectra Compared to AT2017gfo</td>
<td>108</td>
</tr>
<tr>
<td>7.3</td>
<td>ARTIS Spectra</td>
<td>111</td>
</tr>
<tr>
<td>7.4</td>
<td>SEDONA Kilonova Models</td>
<td>114</td>
</tr>
<tr>
<td>7.5</td>
<td>TH13 Analysis of AT2017gfo</td>
<td>116</td>
</tr>
<tr>
<td>7.6</td>
<td>Recombination Rates of Nd</td>
<td>117</td>
</tr>
</tbody>
</table>
**Acronyms**

AGB - Asymptotic Giant Branch

BB - Blackbody

BBH - Binary Black Hole

BHNS - Black Hole Neutron Star

BNS - Binary Neutron Star

CCSN - Core Collapse Supernova

CE - Common Envelope

EB - Energy Balance

EM - Electromagnetic

GRB - Gamma-ray Burst

GW - Gravitational Wave

HMNS - Hypermassive Neutron Star

IR - Infra-red

ISM - Interstellar Medium

JWST - James Webb Space Telescope

KN - Kilonova

LC - Lightcurve

LTE - Local Thermodynamic Equilibrium

LIGO - Laser Interferometer Gravitational Observatory

MC - Monte Carlo

MIR - Mid Infra-red
MRI - Magneto-rotational Instability
NLTE - Non-Local Thermodynamic Equilibrium
NIR - Near Infra-red
NS - Neutron Star
NT - Non-Thermal
PE - Photoexcitation
PI - Photoionisation
RT - Radiative Transfer
SED - Spectral Energy Distribution
SMNS - Supramassive Neutron Star
SN - Supernova
SPH - Smoothed Particle Hydrodynamics
TI - Thermal Ionisation
TLA - Two Level Atom
TOV - Tolman-Oppenheimer-Volkoff
UV - Ultraviolet
ZAMS - Zero-Age Main Sequence
1 Introduction

The question of origin has long been in humanity’s mind, and in particular, that of the Earth and our own origin. While humans are primarily composed of light elements like carbon, hydrogen, and oxygen, heavier elements are still present in smaller quantities, and play key roles in our functioning. Among those, elements heavier than iron are present in even smaller trace amounts, and yet are vital to our existence. So where do all these elements that make up a person come from?

We know that the lightest elements, hydrogen, helium and lithium, come from primordial nucleosynthesis, whether in a Big Bang scenario or otherwise. We also know that subsequent heavier elements up to iron are made by stellar fusion, which has lead to the saying that humans are made of stardust. Now what about the trace elements heavier than iron? These have long been puzzling, and require a different creation process than fusion. At the present day, we believe we may have the answer to this question.

Two processes are suggested for these heavy elements, called the slow and rapid neutron capture processes respectively (s/r-process Burbidge et al. 1957; Cameron 1957). As the name implies, these nuclear processes involve the capture of free neutrons by a seed nucleus, which forms an isotope with a higher atomic mass. This newly formed isotope can then undergo radioactive decay depending on its stability, potentially forming a different element than the original parent species. The main distinction between the two processes is the rate at which the neutron capture occurs. In the s-process, the capture is slower than beta-decay, whilst in the r-process, the capture is faster. This also implies a difference in neutron density required for the two processes, where the r-process requires a much higher neutron density than the s-process. The consequence of captures occurring slower or faster than decays is that not all elements and isotopes are created equally by both processes, with some isotopes even being entirely s or r-process synthesised. However, it is estimated that each process contributes approximately equally to the abundance of
elements heavier than iron in the Universe (see Prantzos et al. 2020; Cowan et al. 2021; Perego et al. 2021; Arcones & Thielemann 2023, for reviews).

The s-process is relatively well known to occur in several sites, such as in low and intermediate mass stars on the asymptotic giant branch (AGB) (Sanders 1967; Schwarzschild & Härm 1967), as well as in the He-burning cores of massive stars (Peters 1968). Sites allowing the r-process to operate were long debated, and to this day, no single scenario is able to fully reproduce the measured abundances of r-process elements in the Solar System (e.g. Sneden et al. 2008; Siegel et al. 2019; Côté et al. 2019a; Arcones & Thielemann 2023). However, a promising site for the generation of many r-process elements comes from the merging of a binary neutron star (BNS), or neutron star (NS) black hole (BH) system, where the dense neutron rich matter ejected by the process allows for rapid neutron capture to occur (Lattimer & Schramm 1974; Symbalisty & Schramm 1982). These merger events are not only a strong source of gravitational wave (GW) radiation, but many of the heavy elements produced during this merger will be unstable isotopes, which decay to power a supernova (SN) like transient called a kilonova (KN) (Li & Paczyński 1998). To this day, only a single KN has been observed both in GW and electromagnetically (EM): the famous AT2017gfo (Abbott et al. 2017a). KNe have been inferred from excess emission in the afterglow of gamma-ray bursts (GRBs) (e.g. Rastinejad et al. 2022; Gillanders et al. 2023b; Levan et al. 2023), but the debate remains active on the conclusive nature of these objects.

KNe evolve in a manner that is conceptually similar to SNe, though with key differences, particularly with respect to timescale and energy deposition. Immediately after the merger, the KN ejecta are dense and cooling by adiabatic expansion. As they are still closely coupled to the radiation, modelling of this phase requires use of radiation hydrodynamics to explore the coupling of matter and radiation. The hydrodynamic effects play out quickly however, on a matter of milliseconds (e.g. Kawaguchi et al. 2021; Neuweiler et al. 2023), after which the ejecta expand homologously and radiative transfer (RT) can be decoupled from hydrodynamics. In the initial homologous expansion phase, the ejecta are still dense enough such that they are optically thick to most wavelengths. This leads to the formation of a photosphere in the ejecta, below which material is optically thick and thus not observable, hence this regime is called the photospheric phase. At later times, after the peak of the bolometric lightcurve, the KN enters a lower density phase, where some wavelengths become more optically thin. As time goes on, this occurs for the
majority of optical and infra-red (IR) wavelengths, and dropping opacity allows the ejecta
to become transparent, which is then referred to as the nebular phase. It is also expected
that the spectra become dominated by atomic emission lines, which provides an excellent
opportunity for the identification of elements created in the KN, and in particular deter-
mination of their masses and distribution geometry, since the emission now probes the
whole ejecta rather than a thin surface layer.

The study of KNe is interesting for various reasons, and provides an opportunity to
answer many different open science questions. Firstly, since NS mergers are theorised
to be one of the major, if not the most important, sources of r-process elements
in the Universe, being able to correctly predict and identify elemental signatures in the
resultant KN spectra is critical. This also implies being able to correctly infer ejecta
masses, velocities, and geometries, as composition is expected to vary depending on these
parameters. Since KNe are born of neutron rich matter from NS, inferring information
from them may also provide hints as to the behaviour of the nuclear matter, such as
the stiffness of the equation of state, which may help constrain nuclear matter models
(Oertel et al. 2017; Bauswein et al. 2020; Foucart et al. 2023). Other properties from
the progenitor systems, such as total mass and spin (e.g. Shibata &Hotokezaka 2019;
Köppel et al. 2019; Salafia et al. 2022), can also help constrain and define stellar evolution
models, especially with respect to binary formation and evolution, as well as core-collapse
SN mechanisms (e.g. Chruslinska et al. 2018; Gompertz et al. 2022; Sgalletta et al. 2023).

The area of NS mergers and KNe is still relatively new, and many aspects are the
subjects of novel and interesting studies, from the nuclear network calculations regarding
nucleosynthesis (e.g. Wanajo et al. 2014; Rosswog et al. 2018; Wanajo 2018; Prantzos
et al. 2020; Nedora et al. 2021; Wu &Banerjee 2022), to the modelling of the late time
emission from the gamma ray burst (GRB) afterglow (e.g. Margutti &Chornock 2021;
James et al. 2022; Gottlieb et al. 2023). The work conducted for this PhD thesis addresses
the nebular phase modelling of KNe. To this end, I modify the SN spectral modelling
code SUMO (Jerkstrand et al. 2011, 2012), and apply it to KN modelling in attempts to
derive conditions in the ejecta, as well as the emergent lightcurve and spectra.

The focus of this PhD thesis has initially been to test two commonly used assumptions
in KN radiative transfer (RT) simulations. The later focus was shifted to the generation
of NLTE spectra of KNe, produced in a fully self-consistent manner. The first publication
included in this work investigates when the KN transitions from the early steady-state neb-
ular phase, to the later time-dependent phase, and the significance of the effects stemming from such a transition. This transition is expected to occur when thermal reprocessing timescales become long compared to the evolutionary time, and may lead to significantly different temperature and ionisation structures compared to the steady-state phase. This will in turn affect the total bolometric luminosity, and spectral output, and is thus important to correctly consider when modelling emission in the nebular phase (see Jerkstrand 2017, for a review in a SN context).

The second publication considers the validity of the local thermodynamic equilibrium (LTE) approximation for the early nebular phase, commonly employed for simple KN models, or those focussing on earlier times after the merger. As the KN expands into the nebular phase, the decreasing density is expected to push the ejecta into the non-local thermodynamic equilibrium (NLTE) regime. However, NLTE modelling is significantly more computationally expensive, and complex to implement compared to LTE models. As such, it is important to know when NLTE computations are necessary for accurate calculations.

The third publication studies the spectra of three, multi-zone, 30-element homogeneous composition models from 5 to 20 days after merger. The compositions and associated decay power were taken from the nuclear network calculations of Wanajo et al. (2014), and so are self-consistent. The density profile was taken from the hydrodynamical simulations of Kawaguchi et al. (2021), representative of a relatively ’standard’ BNS merger scenario. The compositions were parametrised by the electron fraction, leading to light, medium, and heavy composition models. Studying these, it was found that the elements Rb, Sr, Y and Zr play key roles in lanthanide-free ejecta, while the lanthanides Nd, Sm and Dy dominate otherwise. Fluorescence and resonance scattering were also identified as key transfer processes in this time-frame, showing that they effectively transport flux redwards due to extreme line blocking at UV/optical wavelengths. This paper provides the first, fully consistent NLTE spectra of KNe in the literature, and though the ejecta models remain relatively simple, provides a solid foundation for more detailed studies to follow.

In order to understand nebular phase KNe, as well as the significance of the findings in this thesis’ publications, one must first understand several more fundamental parts of NS merger and KN physics. I begin by briefly addressing the progenitor systems of KNe in Chapter 2, as well as the gravitational radiation associated to merging compact. I then move on to the hydrodynamical aspect of the merger, and the associated implications for
r-process nucleosynthesis in Chapter 3. Following this, in Chapter 4 I take a closer look at the electromagnetic emission resulting from the merger, with a focus on the KN transient. I then present in detail the physics relevant to the nebular phase, which is the main object of study in this work, in Chapter 5. The functioning of SUMO for modelling KNe is described in Chapter 6. Following this, a summary on the key operational differences with other currently used RT codes in the field is found in Chapter 7. The papers included in this PhD thesis are summarised in Chapter 8, and the papers themselves can also be found at the end of this thesis. Finally, I offer my views on the current state and future of the field in Chapter 9.
In order for a KN to be produced, a merger consisting of a NS, and either a second NS, or a BH must occur. In this chapter, I briefly review the theorised steps leading up the compact object merger preceding the KN.

2.1 Compact Object Remnants

In both cases of a NS and a BH, a massive star ($M_{ZAMS} \gtrsim 8 \, M_\odot$) must undergo a core collapse SN (CCSN). Right before the core collapse event, these massive stars will have an "onion-like" elemental shell structure, as shown Figure 2.1. As the burning of metals inside the star continues to add mass to the core, the interior density will reach a point where nuclear fusion is unable to sustain the core against gravity, at which point core collapse occurs. This triggers the CCSN, which in a matter of seconds violently expels the star’s outer layers, spreading elements up to iron and nickel back into the interstellar medium (ISM). The exact mechanisms of the core collapse and explosion, such as the role of shocks, neutrinos, magnetic fields etc., are still intensely studied and debated today (see e.g. Janka 2012, for a review). However, the nature of the remnant that is left behind is widely accepted to depend heavily on the mass of the progenitor. Generally speaking, progenitors of masses up to $\sim 25 \, M_\odot$ are thought to leave behind NSs, while heavier progenitors of $\gtrsim 40 \, M_\odot$ are thought to leave behind BHs (Carroll & Ostlie 2006; Janka 2012), though the landscape is complex and actively researched (e.g. Sukhbold et al. 2016; O’Connor et al. 2018; Kuroda et al. 2022).

Supported against gravitational collapse by neutron degeneracy pressure and the strong nuclear force, NSs are expected to exist in a mass range roughly defined by the Chan-
drasekhar (∼ 1.4 $M_\odot$) (e.g. Hawking & Israel 1989) and Tolman-Oppenheimer-Volkoff (TOV) limits (∼ 2.2 $M_\odot$) (e.g. Cho 2018), representing the limits of electron and neutron degeneracy pressures respectively, with additional support from repulsive nuclear forces in the TOV case. Left alone without environmental interaction, NSs are expected to slow down from their initial rotation and cool due to magnetic breaking. However, environmental interactions with the ISM and/or other objects around them may allow a NS to continue evolving in various ways. These interactions are central to the creation of compact object binaries, their eventual coalescence, and the creation of a KN.

### 2.2 Binary Formation and Evolution

The formation of a compact object binary which will merge within the lifetime of the Universe is no trivial matter. Ignoring for now external environmental effects, a compact object binary will lose energy through GWs, causing the orbital separation to decrease, and eventually an inspiral and coalescence to occur. The GW luminosity, or rate of energy loss, is given by (e.g. Peters 1964):

\[
L_{GW} = \frac{16\pi^4 c^6}{15\hbar G^4} \frac{M_p^3}{a^5} \left(\frac{\dot{M}}{M_p}\right)
\]
\[ \dot{E} \propto \frac{m_1^2 m_2^2 (m_1 + m_2)}{a^5} \]  

(2.1)

where \( m_1 \) and \( m_2 \) are the heavier and lighter binary components respectively, and \( a \) is the separation.

While the emission at large separations is exceedingly small, the GW luminosity during the final seconds of a (binary BH) merger can reach approximately \( \sim 10^{57} \text{ erg s}^{-1} \) (e.g. Cardoso et al. 2018), "outshining" all the stellar EM emission in the visible Universe combined! However, the small energy loss rate at large separations implies that binaries have to form relatively close to each other if GWs are solely responsible for the orbital inspiral. For example, two \( M \sim 30 M_\odot \) BHs need an initial separation of \( a_0 \sim 50 R_\odot \) to merge within 14 Gyr (Mandel & Farmer 2018). Using the formula above, this separation drops to approximately \( 1.5 R_\odot \) for two \( 1.5 M_\odot \) neutron stars! The massive stars required to give birth to BHs and NSs expand as they evolve though, and will at some point have radii much larger than several hundred \( R_\odot \). Should two stars be born within this distance, it is likely they will interact with each other in some way beyond GW emission.

The existence of compact binary systems is now a proven fact, and even before the advent of GW observations, were well known (e.g. Tauris & van den Heuvel 2006). Observationally, X-ray binaries provided insight into the evolutionary process of a BH with a less evolved star (see e.g. De Loore & De Greve 1975; Fabbiano 2006), while the famous Hulse-Taylor pulsar demonstrated the existence of binary NS systems (Hulse & Taylor 1975; De Loore et al. 1975), and luminous red novae are now associated with binary evolution (Bond et al. 2003; Ivanova et al. 2013; Howitt et al. 2020). Since the construction of the GW observatories LIGO and Virgo, merging compact objects are now also well established by numerous GW observations (see e.g. Abbott et al. 2023, for a recent summary).

As of today, several theories exist to explain the formation and evolution of binary compact object systems, and how these can merge within the Universe’s lifetime. These can be broadly split into two categories: isolated evolution (e.g. common envelope and chemically homogeneous evolution), and dynamical interactions. I briefly summarise the channels here, focussing mostly on the common envelope evolution which is favoured for BNS systems (see also Mandel & Broekgaarden 2022).
2.2.1 Common Envelope Evolution

Perhaps the most studied evolutionary channel for a tight compact binary is that of common envelope (CE) evolution (see e.g. Livio & Soker 1988; Postnov & Yungelson 2006; Ivanova et al. 2013; Mandel & Broekgaarden 2022). This formation channel is an isolated evolution approach that relies on a tightening of the binary at the right moment by dynamically unstable mass transfer, known as the common envelope phase. The CE channel was initially proposed for BNS systems in the context of the Hulse-Taylor pulsar (Smarr & Blandford 1976). I summarise here the main steps required for this phase to create a tight compact binary:

1. Two massive stars ($M_{\text{ZAMS}} \gtrsim 8 M_\odot$) are born in a binary with moderate separation.

2. The heavier star (primary) reaches the end of the main sequence (MS) and expands. If this expansion is sufficient, it will pass the equipotential surface known as the Roche Lobe, and transfer mass to the lighter companion (secondary). This mass loss may be non-conservative, i.e. the secondary does not accrete all the mass lost by the primary. Combined with the Roche Lobe overflow, mass loss from winds can widen the binary.

3. The primary loses its hydrogen envelope due to mass loss, becoming a Wolf-Rayet star.

4. The system continues to widen by stellar wind mass loss. The primary collapses and becomes a compact object. The SN may eject some mass and yield a natal kick.

5. Sometime later, the secondary reaches the end of the MS, and mass transfer flows towards the primary remnant. This remnant is now less massive than the secondary, so the mass transfer adds specific angular momentum to it. Conservation of angular momentum forces a shrinking of the binary.

6. The more mass is donated, the more the donor overflows the Roche Lobe as the binary hardening is faster than the donor’s reduction in size. The runaway process is dynamically unstable and leads to the formation of a common gas envelope around the binary. The drag from this envelope leads to rapid inspiral, with orbital energy deposited to the envelope.
7. The envelope is eventually expelled, leaving behind a bare binary consisting of a compact object and a Wolf-Rayet star.

8. The secondary continues its evolution until collapse into a compact object, and we now have a BBH, BNS or BHNS system which can (eventually) coalesce by GW emission.

Many aspects of this channel remain quite uncertain (e.g. Iorio et al. 2023). Notably, mass loss from stellar winds, especially the dependency on metallicity and variation during the various evolutionary phases of massive stars is still uncertain (e.g. Vink 2011; Higgins et al. 2023). There are also many other uncertainties, such as the rate of transfer of angular momentum in the system, and the response of a star to mass loss (Pavlovskii et al. 2017; Romagnolo et al. 2023), which is dependent on the star's equation of state. These factors will affect the Roche Lobe overflow process and onset of the common envelope phase, notably with respect to its lifetime (i.e. it can be ejected if it receives too much energy), the degree of binary hardening (e.g. Kruckow et al. 2016; Fragos et al. 2019; Klencki et al. 2021), as well as mass lost during envelope ejection (e.g. Laplace et al. 2020). There is also a large degree of uncertainty tied to the collapse of the massive stars, as the SN may provide large natal kicks that can disrupt the binary (Repetto et al. 2012; Fryer et al. 2012; Mandel 2016). Variations on this formation channel also propose double-core common envelopes when both stars are already evolved (Bethe & Brown 1998; Belczynski et al. 2002; de Mink et al. 2009), or binary hardening by non-conservative mass transfer without an envelope phase (Pavlovskii et al. 2017; van den Heuvel et al. 2017; Cehula & Pejcha 2023). Typical formation rates inferred from this channel vary between $\sim 10 - 100 \text{ Gpc}^{-3} \text{yr}^{-1}$ (Mandel & Farmer 2018; Mandel & Broekgaarden 2022), depending on the assumptions taken related to the aforementioned uncertainties.

### 2.2.2 Chemically Homogeneous Evolution

The main issue with forming a tight compact binary from stars at zero-age main sequence (ZAMS), is that these massive stars will expand at some point during their evolution. This will lead either to a CE phase, or an early stellar merger and thus result in a single remnant as opposed to a binary system. Chemically homogeneous evolution suggests a method to avoid having those massive stars expand during their evolution, thereby allowing tight binaries to form from the beginning (de Mink et al. 2009; de Mink & Mandel 2012).
This also requires that most of the star’s mass be converted into a compact object, typically a BH, as opposed to the large fraction which typically becomes the helium core. The steps for this channel are outlined here:

1. Close binary companions raise tides on each other, with tidal energy dissipation leading to tidal locking of the two stars. This means the stars have synchronised rotations, and are spinning to a few tens of percent of their break up velocity, i.e. very fast rotation.

2. The rapid rotation generates important temperature gradients between the poles and equators (Eddington 1925; Sweet 1950), leading to large scale circulation in the stars, and important chemical mixing (Endal & Sofia 1978; Heger et al. 2000; Yoon et al. 2006; Szécsi et al. 2015). This may occur to the extent that hydrogen is transported to the core, and helium outwards until the star is practically all helium.

3. The stars are essentially Wolf-Rayet stars, and contract at the end of the MS. If metallicity is low, stellar wind mass loss is small and the binary does not widen, mass transfer can be avoided.

4. The stars collapse to compact objects, and a compact object binary is formed. This channel, as unlikely as it seems given the evolutionary fine tuning required, could potentially present a viable formation channel for the most massive GW signals, and thus would typically form BBH systems without any NS. It may however be a more reasonable scenario for binaries formed in the early Universe from massive, metal poor Population III stars (Marigo et al. 2001; Mandel & Farmer 2018).

2.2.3 Dynamical Interactions

In this channel, the compact objects do not start out as binaries, but find each other through various interactions with other objects, typically in a dense stellar environment such as globular clusters. Possible steps for this formation channel are as follows:

1. Two compact objects form in a dense stellar environment. Particularly for BHs, they move towards the center by mass segregation, being typically heavier than the other cluster objects (Spitzer 1969).
2. Three-body interactions then allow the formation of a binary system, in which the third object receives enough kinetic energy to leave the system. This is typically the lightest object which is ejected in favour of the two heavier objects (Hills & Fullerton 1980).

3. Should the binary be hard, i.e. the speed of the components are higher than the typical speed of the cluster’s stars, then interactions with nearby stars will gradually tighten the binary (Heggie 1975). Soft binaries will however likely be disrupted.

4. In a dense enough system, the rate of interactions with third objects will be sufficiently high to tighten the binary to the point where GW emission will allow merging within the age of the Universe.

Dynamical interactions do not necessarily have to proceed in the exact fashion described above. The exact steps undertaken may be somewhat different in order to form a BNS system, where the NS are not the heaviest objects in the cluster. This channel is typically disfavoured for BNS systems, since their lighter masses will typically result in them being ejected during interactions (Mandel & Broekgaarden 2022).

2.3 Gravitational Wave Signals

All compact object binaries emit GWs as they orbit each other. At the initially large separations of the binary components, the frequencies emitted are too low to be observed by the current GW observatories LIGO (Laser Interferometer Gravitational-Wave Observatory) and Virgo (see Figure 2.2). As the binary separation shrinks, the signal amplitude and frequency increase until these are high enough to be detectable. LIGO is technically capable of detecting signals over a wide range of frequencies from 10 Hz to 10 kHz (Martynov et al. 2016).

The amplitude and frequency increase accelerate with time, reaching maxima at the moment of coalescence. This forms a characteristic ‘chirp’ signal that is common to all merging compact object binaries. Right after the moment of coalescence, a ‘ringdown’ signal is emitted (see e.g. Schmidt 2020, for a review on waveforms). Figure 2.3 shows the waveform of GW170817, the first binary neutron star merger ever observed in gravitational radiation (Abbott et al. 2017b).
Figure 2.2: Schematic illustrating the function of LIGO when detecting GWs from a merging BH system. Credit: LIGO collaboration. Taken from https://phys.org/news/2019-05-ligo-virgo-neutron-star-smash-ups.html.
Figure 2.3: The GW signal (and GRB emission) of GW170817, the first ever binary neutron star merger observed by the LIGO interferometer. The characteristic 'chirp' is well visualised in frequency space. To this day, this remains the only GW detected BNS merger that has also been successfully observed in an electromagnetic follow-up. Taken from Abbott et al. (2017b).
The exact properties and shape of the waveform depend on many factors, which makes it possible to extract information about the binary system, although this is challenging due to parameter degeneracy. The analysis is typically done by Bayesian inference, with observatories like LIGO having sophisticated model matching methods in order to find the best fitting template waveform to the observed data. A key parameter for triggering follow up EM searches is the component masses. If at least one component should lie in the mass range typical for a NS, a KN may have been created, particularly if both components are well constrained to NS masses. As this PhD thesis is about KN ejecta and emission, I focus here on the mass determination from GW signals.

The component masses are typically derived from various, easier to measure quantities, such as the total mass \( M = m_1 + m_2 \), as well as a quantity called the ‘chirp’ mass:

\[
M_C \equiv m_1^{3/5} m_2^{3/5} M^{-1/5}
\]  

(2.2)

The ringdown frequency after coalescence is a function of total mass, and is typically better measured for heavier systems, i.e. for heavy BBH systems. The chirp mass determines the frequency evolution during the inspiral before coalescence, and is typically better measured for lower mass binaries (Mandel & Farmer 2018). Measurement of these two quantities allows the individual component masses to be derived. Should \( M \) or \( M_C \) be poorly constrained, an additional parameter \( q = m_2/m_1 \leq 1 \), called the mass ratio, can be estimated. However, this is degenerate with compact object spins (Poisson & Will 1995; Hannam et al. 2013).

Mass is not the only parameter that can be inferred from GW observations. Notably, compact object spins can be estimated, though this is a difficult endeavour as there is degeneracy with binary eccentricity (e.g. Schmidt 2020). Aside from the intrinsic properties of the binary, location and distance can also be gauged from GW observations (e.g. Veitch et al. 2015; Ashton et al. 2019), though this can be difficult due to dependencies on orientation and emitted signal strength (e.g. a BBH merger signal has a larger amplitude than a BNS merger signal, and is detectable to further out). Localisation is typically done by triangulation, and so requires several observatories to detect the GW signal.

During the observing runs thus far, LIGO-Virgo have managed to observe many compact object binaries, with several possible BNS or BHNS mergers which have yielded no EM follow-up detections, either due to poor sky localisation or large distances. Thus far, only a single NS merger has been successfully observed in both GW and EM emission:
Figure 2.4: Compact object masses as observed electromagnetically and observed by LIGO via gravitational waves at the end of the third observing run O3. The colours refer to the nature of the object, and are described at the top of the figure. The 3 - 5 $M_\odot$ "mass gap", is a sparsely populated area considered extremely light for BHs, but too heavy for NSs. Credit: LIGO -Virgo / Aaron Geller / Northwestern. https://www.ligo.caltech.edu/image/ligo20211107a

the famous GW170817/AT2017gfo event (Abbott et al. 2017a). GW observations from LIGO-Virgo have, however, allowed the identification of many otherwise unknown compact objects, especially for stellar mass BHs in previously sparsely populated parts of parameter space (see Figure 2.4). However, since there is a bias towards detecting massive systems with strong GW emission, most of the LIGO-Virgo observations have been of heavy BBH mergers (Abbott et al. 2021, 2023).

Considering the GW170817/AT2017gfo event in terms of gravitational radiation, this event was extremely fortunate, in that it was relatively nearby at $40 \pm 8$ Mpc, and had a favourable orientation such that the GW sky localisation was relatively small, only $31$ deg$^2$ (see Figure 2.5). It was confirmed as a BNS merger as both component masses were inferred to be $\sim 1.2 - 1.6$ $M_\odot$, too light for BHs. For comparison, the second BNS merger observed by LIGO in 2019, with the GW signal named GW190425, was much further away at $156 \pm 41$ Mpc, and with a broader sky localisation of $7460$ deg$^2$ (Hosseinzadeh et al. 2019). As such, even with a thorough follow up EM campaign, this BNS merger was never seen electromagnetically, and neither has any GW detected BNS merger since. The component masses of this merger were poorly constrained due to parameter inference
Deviations with the NS spins, with some estimates yielding a primary component mass as high as $\sim 2.5 M_\odot$ (Abbott et al. 2020). This lead to some debate as to whether the event was actually a BHNS merger, with an expected fainter associated KN (e.g. Kyutoku et al. 2020).

These difficulties with respect to localisation, as well as the bias towards strong GW signals from heavy binaries, provide a potential explanation as to why, thus far, only one EM counterpart to a GW detected BNS or BHNS merger has been observed (BBH mergers should produce no light), whereas several more NS mergers were observed in GW emission. Depending on the distance and angle of the binary compared to the GW observatories, the sky localisation area yielded by LIGO-Virgo may be quite large, and provide little hope for successful follow up EM localisation. As such, EM transient observation facilities, are required to fully localise the transient and provide a smaller search region for subsequent EM observations. While GRB facilities may provide excellent localisation if the GRB is beamed at a convenient angle, the localisation may be worse than for the GW emission if this is not the case, as seen in Figure 2.5. In this case, transient sky surveys in optical wavebands are key to detecting the GRB afterglow, or the KN itself, as was the case for
AT2017gfo (Abbott et al. 2017a). This is also a challenge however, given the rapid fading of EM transients associated to BNS and BHNS mergers (see Chapter 4).

While GW signals can yield a lot of information about the nature of the merging compact binary, they are only able to tell part of the story. Electromagnetic emission associated to the merger, such as the GRBs and KNe also provide significant information that cannot be communicated by GW emission, aside from aiding in localisation. KNe in particular are the key observable with respect to the creation of heavy r-process elements and their radioactive decay. In the next Chapter, I describe how the r-process occurs in the ejecta, and the associated energy generated by radioactive decay which powers the KN.
3 Ejecta and r-process Nucleosynthesis

As the BNS or BHNS system merges, neutron rich material will be violently ejected by several different processes, providing an ideal site for r-process nucleosynthesis. Before delving into the details of ejecta hydrodynamics relevant to the papers in this thesis however, it is useful to understand some of the history and physical context behind the importance of the r-process.

3.1 Basics of the Rapid Neutron Capture Process

As previously described, the r-process is the nucleosynthesis of elements by rapid neutron capture. Since the r-process requires the capture of neutrons on very short timescales, a correspondingly large free neutron density is required. While the s-process is typically attainable at densities of $n_n \sim 10^6$ cm$^{-3}$, the r-process requires a much higher density of $n_n \gtrsim 10^{24}$ cm$^{-3}$ (Prantzos et al. 2020). The degree of neutron richness is quantified by the electron fraction:

$$Y_e \equiv \frac{n_p}{n_n + n_p}$$  \hspace{1cm} (3.1)

where $n_n$ and $n_p$ are the number densities of protons and neutrons, both bound and free, respectively. The yield of the r-process nucleosynthesis is heavily dependent on $Y_e$, with the heaviest of elements only being formed at extremely low electron fractions $Y_e \lesssim 0.2$ (more on this below).

The details of the r-process have long posed an intriguing challenge from both astrophysical and nuclear physics perspectives. Since the r-process creates isotopes close to the neutron drip line, many quantities, such as nuclear binding energies, $\beta$-decay rates, fission fragment distributions etc. are not known experimentally, limiting the accuracy of nuclear network calculations modelling the r-process. Furthermore, astrophysical sites
providing suitable conditions for r-process nucleosynthesis remained elusive for a long time. Since the establishment of the r-process as a necessity for the creation of heavy elements (Burbidge et al. 1957; Cameron 1957), many theories have been presented in attempts to explain the origin of r-process elements in the Universe. Neutrino driven winds from proto-neutron stars emerging seconds after a CCSN were initially suggested (Duncan et al. 1986; Woosley et al. 1994; Qian & Woosley 1996), but were later excluded as dominant sites due to the necessity of low electron fractions and high entropy (Arcones et al. 2007; Martínez-Pinedo et al. 2012), though they may still produce heavy, proton-rich nuclei (e.g. Wanajo 2023). CCSNe themselves have also been suggested as low yield r-process sites (Mathews et al. 1992; Sneden et al. 2008), and rare types of SNe (e.g. collapsars Woosley 1993; MacFadyen & Woosley 1999; Siegel et al. 2019) continue to be suggested today (Côté et al. 2019a; Cowan et al. 2021; Arcones & Thielemann 2023).

The discovery of the first binary pulsar (Hulse & Taylor 1975), alongside the first theoretical propositions of compact object binaries (Lattimer & Schramm 1974, 1976) paved the way for NS mergers to emerge as potential r-process sites several years later (Symbalisty & Schramm 1982). Landmark theoretical papers predicting the mass yields, as well as a connection to short gamma ray bursts (sGRBs) followed these first papers (e.g. Paczynski 1986; Eichler et al. 1989; Freiburghaus et al. 1999), with initial simulations of BNS mergers broadly reproducing solar r-process abundances (Rosswog et al. 1999). The first EM predictions of what the associated transients should look like, and the first suggestion for the name of kilonova occurred a decade later (Metzger et al. 2010). However, observations of BNS mergers remained difficult until the construction of GW observatories such as LIGO-Virgo, which finally allowed a KN to be successfully observed. Today, compact binary mergers are commonly accepted to be a promising source of r-process elements (see e.g. Perego et al. 2021; Arcones & Thielemann 2023, for reviews).

In the past decade, much work has been carried out on the r-process, with a particular focus on conditions within NS merger ejecta. A key objective is to determine what conditions are required to reproduce the solar abundance pattern of r-process elements (e.g. Rosswog et al. 2018; Pritschenko 2019; Prantzos et al. 2020; Nedora et al. 2021). The r-process synthesises elements roughly from Ga (Z = 31, A ∼ 70) all the way to U (Z = 92, A ∼ 238), though naturally, not all elements and isotopes are produced in equal amounts. Notably, the degree of neutron richness of the material, quantified by the electron fraction given in Equation 3.1, will affect the elements created, with relative
abundances also depending on neutron capture cross-sections and $\beta$-decay rates.

The solar abundance pattern of elements is shown in Figure 3.1. The r-process generally produces three main abundance peaks around mass numbers $A = 80, 127,$ and $195$, with the lanthanides ($138 \leq A \leq 173$) maximised around $A \sim 165$. The three main abundance peaks occur due to the build up of nuclei with closed neutron shells, given by the ‘magic numbers’ $N = 50, 82, 126$ (e.g. Arcones & Thielemann 2023). There is still ongoing work determining the exact fraction of elements stemming from the r-process or s-process, and the exact yields vary between studies and methods used (see e.g. Sneden et al. 2008; Pritychenko 2019; Prantzos et al. 2020; Lodders 2021). It is generally agreed upon however, that lower $Y_e$ media will synthesise heavier elements (see Figure 3.2).

Although NS mergers are an excellent site for r-process nucleosynthesis, it is probable that they are not the sole astrophysical sites in the Universe. The question is presently whether they are the dominant source of r-process elements. The required production rate of heavy ($A > 140$) r-process elements needed to account for the abundances inferred for the Milky Way is found to be approximately $2 \times 10^{-7} \, M_\odot \, \text{yr}^{-1}$ (Qian 2000; Qian}
Figure 3.2: Plots of mass fraction vs. mass number for an ejecta with $v_{ej} = 0.25$ c, and entropy $s = 15$ k$B$, showing the effect of electron fraction $Y_e$ on r-process yields, in red. The blue points are solar system r-process abundances. Taken from Rosswog et al. (2018).
The initial BNS merger rate densities inferred from the first LIGO observation were very broad, \( R_{\text{BNS}} = 320 - 4740 \text{ Gpc}^{-3} \text{ yr}^{-1} \). Combined with r-process abundances inferred for the Milky Way, these rates implied a galactic r-process mass yield per event of roughly (Metzger 2019):

\[
< M_r > \approx 10^{-2} \, M_\odot \left( \frac{R_{\text{BNS}}}{1250 \, \text{ Gpc}^{-3} \, \text{ yr}^{-1}} \right)^{-1}
\]  

(3.2)

In comparison, the estimated mass yields of AT2017gfo ranged from \( \sim 0.02 - 0.08 \, M_\odot \) (Metzger 2019).

However, the lack of further BNS merger detections has led to a broadening in merger rate estimates compared to the initial LIGO value, to \( 10 - 1700 \, \text{ Gpc}^{-3} \text{ yr}^{-1} \) (Mandel & Broekgaarden 2022), with estimates from other detection methods such as GRBs setting upper limits closer to \( R_{\text{BNS}} \leq 900 \, \text{ Gpc}^{-3} \text{ yr}^{-1} \) (Andreoni et al. 2021; Mandel & Broekgaarden 2022). Accordingly, the galactic r-process mass yield per event varies greatly depending on ejecta masses and assumed merger rates (see e.g. Rosswog & Korobkin 2022, for a review).

BNS mergers may be the main source of r-process elements in the Universe, under certain assumptions (see e.g. Côté et al. 2018; Wanajo et al. 2021), notably with respect to potentially problematic factors such as SN kicks ejecting NS from binaries, or the short delay times required between stellar formation and merger needed to explain r-process abundances in the low metallicity Galactic halo. Lines of evidence from various other considerations, ranging from the abundance of \(^{244}\text{Pu}\) on the sea floor (Hotokezaka et al. 2015), to the actinide abundance of the primordial solar system (Côté et al. 2019b), also suggest rare, high r-process yield events (such as, but not limited to NS mergers) as the dominating source, as opposed to more frequent, low yield processes such as SNe.

However, other studies now imply that various alternative processes play larger role than previously inferred (see e.g. Côté et al. 2019a; Thielemann et al. 2020; Tarumi et al. 2021; Farouqi et al. 2022). These may include faster channels such as magneto-rotational jet SNe (Mösta et al. 2020; Reichert et al. 2021), and/or collapsars (Siegel et al. 2022; Barnes & Duffell 2023). In general, it is suggested that rare SN types may produce abundances relevant to weak r-process patterns observed in low metallicity stars, though the necessity for at least one site with strong r-process abundance patterns, such as NS mergers, remains in order to explain the solar system abundance pattern (see e.g. Perego et al. 2021; Arcones & Thielemann 2023, for reviews).
3.2 Properties of Compact Binary Merger Ejecta

The matter ejected from a compact object binary will vary in its properties, such as mass, velocity, morphology, and composition. Progenitor and merger characteristics, such as the nature of the compact objects merging (e.g. BNS or BHNS), their masses, their spins, and the equation of state describing the NS material, as well as the presence and intensity of magnetic fields will affect the ejecta properties. Generally speaking, ejecta are broadly categorised as dynamical or disc ejecta (see e.g. Metzger 2019; Shibata & Hotokezaka 2019; Radice et al. 2020; Cowan et al. 2021; Perego et al. 2021, for reviews).

Dynamical ejecta are expelled on a timescale of milliseconds, from tidal forces and by the intense compression heating generated at the surface of the two merging bodies (the latter is only present for BNS mergers). They are typically fast moving, of relatively low density, and neutron rich, such that the low electron fraction provides ideal conditions for the nucleosynthesis of the heaviest r-process elements. I discuss the dynamical ejecta properties in detail in Section 3.2.2. Other debris from the merger which is not tidally ejected, or compressed into the remnant compact object, can circularise and form an accretion disc around the remnant. This disc also has outflows, typically with slower velocities, but may dominate ejecta mass over the dynamical ejecta. Should a long lived NS remnant be present, it will provide an additional neutrino wind component which interacts with the disc. The ejecta stemming from the disc and remnant are discussed in Section 3.2.3. A simple graphical representation of the various ejecta types can be viewed in Figure 3.3.

In the case of a BHNS merger, the conditions for having significant ejecta masses have been found to be highly dependent on the spin and mass of the BH. Notably, a high mass BH with low or retrograde spin is expected to more or less "swallow" the NS whole, thus yielding little to no ejecta. Conversely, a smaller BH with very high prograde spin can tidally disrupt the NS before the merger, thus allowing significant ejection of matter (Foucart et al. 2018; Krüger & Foucart 2020). As for BNS mergers, an accretion disc may also form after BHNS mergers and provide disc wind ejecta (Fernández et al. 2020).

Aside from these two main categories of ejecta, other rarer sources may also be of interest. Differentially rotating NSs can produce mass loss during angular momentum redistribution (Fujibayashi et al. 2018; Radice et al. 2018b), though this process depends sensitively on the operation of viscosity, which is poorly understood in this context. Mass
loss can also occur from a long-lived NS remnant as it undergoes Kelvin-Helmholtz contractions and neutrino cooling (Dessart et al. 2009; Perego et al. 2014). The neutrino driven winds from the latter process interacts with the disc and can have important effects on electron fraction. Should the remnant be a magnetar with a powerful magnetic field \( B \sim 10^{14} \text{ to } 10^{15} \text{ G} \) (Thompson et al. 2004), the rapid spinning of the remnant, combined with open magnetic field lines can substantially increase the mass loss by this channel (Mösta et al. 2020; Shibata & Hotokezaka 2019; Kawaguchi et al. 2022). Some studies argue that magnetar driven winds may in fact be responsible for some of the ejecta components of the GW170817/AT207gfo event (Metzger et al. 2018; Ciolfi & Kallinani 2020), though this is not unanimously agreed upon (see e.g. Kawaguchi et al. 2023).

Although the work in this thesis was conducted under the assumption of spherical symmetry, understanding the various ejecta types of which the KN consists, and their properties remains a central part of these studies. The ejecta mass and velocity grid from Paper I and II \( (M_{ej} = 0.01 - 0.1 M_{\odot}, v_{ej} = 0.05 - 0.2c) \) was chosen to cover a broad range of ejecta parameters, corresponding to both dynamical and disc ejecta. Paper III used homogeneous composition 5-zone models, with a total ejecta mass of \( 0.05 \ M_{\odot} \), velocity range of \( 0.05 - 0.3c \) in uniform steps of \( 0.05c \), and following a \( \rho \propto v^{-4} \) density profile. In general, composition will vary radially and by viewing angle, such that homogeneous
compositions are not fully realistic. Knowledge of how ejecta composition varies is also relevant to understanding the emergent spectrum, as it may bear implications for derived quantities such as expansion opacity (see Section 4.2.1), which is addressed in Paper II.

### 3.2.1 Progenitor Properties and Merger Remnants

In the case of either a BNS or BHNS merger, the properties of the ejected material depend strongly on the characteristics of the progenitor system, and in the case of a BNS merger, also on the nature of the remnant left behind (see Figure 3.3); a BHNS merger can only yield a BH remnant. A key characteristic is the total mass of the remnant $M_{\text{tot}}$ (see e.g. Shibata & Taniguchi 2006; Shibata & Hotokezaka 2019). Should this mass be above a critical value of $M_{\text{crit}} \sim 2.6 - 3.9 \, M_\odot$, the remnant is expected to collapse promptly into a BH (Sekiguchi et al. 2011; Bauswein et al. 2013a, 2021), thereby greatly reducing the amount of ejecta. The exact mass limit varies due to the different possible equations of state describing the neutron rich material of the NS (Hotokezaka et al. 2011; Oertel et al. 2017; Bauswein et al. 2020), as well as the compactness and rotation of the NS, and to some extent the mass ratio of the binary ($q = m_1/m_2 \leq 1$) (Radice et al. 2018a; Kiuchi et al. 2019; Köppel et al. 2019).

If the remnant is below the critical mass, it will not collapse immediately into a BH, but rather results in another NS, the exact type of which is determined by several factors. Typically, the remnant NS will be rapidly spinning, with a slightly smaller mass than the total binary mass, as some mass forms the accretion disc, some is ejected, and some is lost to GW and neutrino emission. The rapid rotation of the remnant provides additional (temporary) support against gravitational collapse into a BH alongside the thermal and neutron degeneracy pressures. The mass of the rapidly spinning NS will be above the TOV limit for a non-rotating NS, but remains below the critical mass (Hotokezaka et al. 2013a; Kaplan et al. 2014; Radice et al. 2018a). The remnant for the observed GW170817A BNS merger is thought to have collapsed to a BH relatively rapidly, though it is not completely agreed upon whether the collapse was prompt, or a short lived NS remnant was briefly present (see e.g. Margalit & Metzger 2019; Cowan et al. 2021; Kawaguchi et al. 2023; Just et al. 2023).

The nomenclature of this remnant depends on the type of rotation supporting it against collapse. A NS supported entirely by differential rotation is referred to as a hypermassive
NS (HMNS), while a less massive NS with solid body rotation is a supramassive NS (SMNS). Typically, HMNS are expected to be short-lived, with a lifetime of tens to hundreds of milliseconds, before collapse to a BH occurs (Shibata & Taniguchi 2006; Siegel et al. 2013; Hanauske et al. 2017). SMNS however, are expected to have longer lifetimes, even up to several seconds, as they require global angular momentum loss in order to spin-down to the point where rotational support is no longer sufficient to prevent collapse (Radice et al. 2018a; Beniamini & Lu 2021). The process associated to the spin-down is typically magnetic dipole radiation. There may also exist a case where a particularly low mass binary merger yields an indefinitely stable NS remnant, though this is expected to be extremely rare, if ever occurring at all (Margalit & Metzger 2019). A long lived NS remnant will eject mass by neutrino driven winds, which will interact with the disc, an aspect that is explored more below.

3.2.2 Dynamical Ejecta

Dynamical ejecta are unbound from the merger on the timescale of milliseconds. The exact amount of matter ejected this way depends on the total binary mass, as well as the mass ratio between the two compact objects, the NS spins (Dietrich et al. 2017; East et al. 2019), and the exact equation of state describing the neutron rich material. Typically, heavier, more asymmetric binaries will yield larger amounts of dynamical ejecta due to tidal effects, though the exact amount depends on the above factors in a complex way (see e.g. Kawaguchi et al. 2023). With all these considerations, the mass ejected by dynamical processes is estimated to range between $10^{-4} - 10^{-2} M_\odot$ for BNS mergers (Radice et al. 2016; Fujibayashi et al. 2018; Bernuzzi et al. 2020; Cowan et al. 2021), though can be up to 0.1 $M_\odot$ for BHNS mergers (Foucart et al. 2017; Hayashi et al. 2021; Cowan et al. 2021), should the BH be low mass with high prograde spin as described above. Generally speaking, binary components with high prograde spin tend to increase the ejecta mass, though NSs in BNS mergers are expected to have small spins compared to the orbital frequency of their circularised orbit at times right before coalescence (Rosswog 2015).

Dynamical ejecta are further split into two components originating from two different processes. The first is tidal ejecta arising from tidal interactions during the merger, often forming spiral arms around the equatorial plane. The matter in this ejecta is initially at relatively cool temperatures and is thought to be very neutron rich with low electron
fractions, $Y_e \lesssim 0.15$ (e.g. Perego et al. 2021). This tidal component is typically more important for binaries with small mass ratios $q = m_1/m_2 \ll 1$, i.e. component masses are significantly different (Bauswein et al. 2013b; Hotokezaka et al. 2013b; Ciolfi et al. 2017). For BHNS mergers, this tidal component is expected to be the sole dynamical ejecta component, and thus BHNS systems with small mass ratios and high BH spin will eject larger quantities of matter (Kawaguchi et al. 2015; Kyutoku et al. 2015, 2018). Tidal ejecta from a BHNS merger is expected to be much more concentrated around the equatorial plane than for a BNS merger, and may only have one tidal tail. It may be even more neutron rich than in the case of a BNS merger (Kyutoku et al. 2018), where neutrino emission may increase the electron fraction.

The second type of dynamical ejecta occurs from the shock heated matter at the contact surface of merging NS (and thus is absent for BHNS mergers), and is expelled by quasi-radial pulsations of the remnant (Bauswein et al. 2013b; Fujibayashi et al. 2018). This matter is shock heated to high temperatures, and due to the quasi-radial nature of the pulsations, is ejected in a broad range of directions (Bauswein et al. 2013b; Hotokezaka et al. 2013b; Radice et al. 2018b). The high temperature of this ejecta drives positron captures, and thus increases the electron fraction $Y_e$ of the ejecta. This can further increase due to absorption of neutrinos from a central remnant (Wanajo et al. 2014; Radice et al. 2016; Ciolfi et al. 2017; Cowan et al. 2021), and can thus lead to relatively high $Y_e$ fractions in the polar regions as neutrino emission is favoured in that direction (Radice et al. 2018b; Shibata & Hotokezaka 2019; Kawaguchi et al. 2023; Just et al. 2023). The shock heated ejecta is typically ejected at velocities higher than the tidal ejecta, peaking around $0.2 - 0.3c$, but with smaller parts capable of reaching $0.6 - 0.8c$ (Shibata & Hotokezaka 2019; Radice et al. 2020; Perego et al. 2021).

In terms of composition, dynamically ejected material is expected to have a broad range of electron fractions, with simulations predicting an electron fraction in the range of $Y_e \sim 0.1 - 0.4$ (Radice et al. 2016; Tanvir et al. 2017; Metzger 2019; Kawaguchi et al. 2020; Just et al. 2023), depending on merger remnant and neutrino transport scheme. This range of electron fraction allows for the nucleosynthesis of many r-process elements. Higher electron fractions closer to $Y_e \sim 0.3$ are required in order to synthesise elements around the first r-process abundance peak at $A \sim 100$ (Wanajo et al. 2014; Rosswog et al. 2018), whilst lower fractions of $Y_e \lesssim 0.2$ are needed in order to synthesis heavier elements at the second and third r-process peaks, around $A \sim 130$ and $A \sim 195$ respectively.
As such, it is generally agreed that both light and heavy r-process elements can be synthesised in dynamical ejecta. Creation of the heaviest r-process elements will occur in the lowest $Y_e$ ejecta, which will typically be the tidal tails, or shock heated ejecta in the event of prompt collapse to a BH.

### 3.2.3 Disc Ejecta

Let us consider first a BHNS merger, or a BNS merger with the remnant promptly collapsing to a BH. During the merger, some ejected material remains bound, and forms an accretion disc around the central remnant, with mass in the range of $M_{\text{disc}} \sim 10^{-3} - 0.3 M_\odot$ (Rosswog et al. 1999; Shibata & Taniguchi 2006; Perego et al. 2014; Radice et al. 2018b; Fujibayashi et al. 2020; Perego et al. 2021). While this is expected to occur for all BNS mergers, a BHNS merger remnant will only have an accretion disc if sufficient material is tidally disrupted outside the BH horizon. The evolution of the disc is determined by hydrodynamical, magnetic, and neutrino driven processes, each with different timescales and impacts.

Initially, the disc has a high mass accretion rate, and generates large amounts of thermal neutrinos which drive matter from the disc in the form of neutrino driven winds (Perego et al. 2014; Hossein Nouri et al. 2018; Fujibayashi et al. 2020). This process occurs on timescales of milliseconds, and typically unbinds only $\sim 1% M_{\text{disc}}$, at a velocity of $v_{ej} \lesssim 0.1c$ (Perego et al. 2014, 2017; Fujibayashi et al. 2020; Perego et al. 2021). Viscosity driven ejection mechanisms operating on the accretion timescale of $\sim 0.5$ s can unbind larger amounts of matter on the order of $\sim 0.1 - 0.4 M_{\text{disc}}$ (Lippuner et al. 2017; Perego et al. 2017; Cowan et al. 2021; Perego et al. 2021). These two processes will always be present in the accretion disc, regardless of central remnant. In the case of prompt BH collapse, the electron fraction of the disc will typically be in a similar range as the dynamical component (Metzger 2019; Shibata & Hotokezaka 2019; Cowan et al. 2021; Perego et al. 2021).

The picture can change significantly if a long lived central NS remnant is produced in a BNS merger. In this case, an additional ejecta component in the form of neutrino driven winds from the remnant will be produced alongside the disc neutrino wind, emitted mostly in the polar direction (Rosswog et al. 2014; Perego et al. 2014; Martin et al. 2015). This additional neutrino wind component from the NS remnant increases ejecta mass and
electron fraction, typically to $M_{ej} \sim 0.05 \, M_{\text{disc}}$, and $Y_e \gtrsim 0.25$, should the NS remnant survive for $t \gtrsim 1 \, \text{s}$ (Perego et al. 2014; Kaplan et al. 2014; Lippuner et al. 2017; Fujibayashi et al. 2018; Cowan et al. 2021; Perego et al. 2021).

Two additional processes stemming from the remnant, and lasting during its lifetime, can also play a large role in disc’s ejecta. Spiral wave winds arising from the oscillation of the NS remnant may push out matter from the disc’s edge (Nedora et al. 2019; Perego et al. 2021). Should the remnant be a magnetar, large scale ordered magnetic fields developed inside the remnant provide magnetic pressure which can also enhance disc ejecta, as long the magnetic fields are present, in the form of magnetically driven disc winds (Metzger et al. 2018). These two processes can provide ejecta masses on the same order as the viscous driven ejecta, e.g. $0.1 - 0.4 \, M_{\text{disc}}$, though at slightly higher velocities $v_{ej} \sim 0.1 - 0.2c$ (Perego et al. 2021).

Regardless of their origin, outflows from the disc will typically produce slower moving ejecta behind the fast moving dynamical ejecta. As such, the disc ejecta can be hidden behind a ‘curtain’ of dynamical ejecta. The viewing angle dependency of this curtain depends on the relative importance of tidal and shock heated dynamical components. In the extreme case of only tidal dynamical ejecta, the disc wind ejecta should be clearly seen from polar angles yet hidden in the equatorial plane. However, should shock heated dynamical ejecta be significantly present, their more isotropic nature may effectively hide the disc wind ejecta from many viewing angles. Viewing angles of the KN thus play a large role in the emergent lightcurve and spectral observations (Bulla 2019; Collins et al. 2023).

### 3.3 Energy Deposition

Many of the r-process isotopes synthesised in the ejecta are unstable. As they undergo radioactive decay, they provide power to the ejecta, and power a quasi-isotropic EM emission known as a kilonova (Li & Paczyński 1998). Without this decay, the transient would be extremely faint as the initial energy from the merger is lost due to adiabatic expansion of the ejecta. Although energy deposition can also occur from other sources, such as central engine power from a long lived NS remnant (see e.g. Yu et al. 2013; Metzger et al. 2018; Wollaeger et al. 2019), or fall-back accretion onto a central BH (see e.g. Rosswog 2007; Foucart et al. 2015; Kyutoku et al. 2015), these are not considered in
the papers included in this thesis, and so I omit their description here.

The decay of unstable r-process isotopes can create internal energy by heating, ionising and exciting the ejecta, and can also do work on the medium by adiabatic expansion. In the studies of this doctoral thesis, the exact fraction of energy going into those channels is found by solving the Spencer-Fano equation (Spencer & Fano 1954; Kozma & Fransson 1992), with the vast majority (∼ 99%) of the energy typically going into heating, as long as the ejecta are singly or a few times ionized. This radioactive heating is the focus of the following section.

Though the KN is radioactively powered in a similar fashion to a SN, a key difference lies in the isotopes which are decaying. While SNe rely on one or a few isotopes of cobalt and nickel to provide power, a KN’s energy source stems from the ensemble decay of many heavy, unstable r-process isotopes, and thus typically yields power laws as opposed to exponential decays (e.g. Metzger et al. 2010; Barnes et al. 2016; Hotokezaka et al. 2017; Kasen & Barnes 2019; Waxman et al. 2019). The primary decay channels are by \( \beta^- \)-decay, which provides electrons/positrons, as well as \( \gamma \)-rays, and neutrinos (which typically escape the ejecta). \( \alpha \)-decay and spontaneous fission may play a role in the lowest \( Y_e \) ejecta, and may provide significant energy due to their effective thermalisation in the ejecta, as shown in Figure 3.4.

As described in Section 3.1, the neutron rich ejecta associated to a BNS merger provide an ideal site for the r-process to synthesise the unstable isotopes. The form of radioactive energy deposition used in Papers I and II is described fully in Paper I. Paper III used the same thermalisation formulae as the first two papers, but with the raw radioactive energy coming from the nuclear network calculations of Wanajo et al. (2014). As such, this section is intended as a general summary of common usage in the community today.
Figure 3.4: Thermalisation efficiencies for radioactive decay products as a function of time. The effect of the magnetic field topology is shown by the different lines: the solid lines are for toroidal field, dotted for radial, and dashed for tangled. This highlights the role of magnetic field trapping in thermalisation efficiency. The ejecta here has $M_{ej} = 0.05$ $M_\odot$ and $v_{ej} = 0.2$ c. Taken from Barnes et al. (2016).

3.3.1 $\beta$-decay

The total energy deposited to the ejecta from radioactivity will be an integral over the plasma losses of all decay products (e.g. Barnes et al. 2016; Kasen & Barnes 2019; Waxman et al. 2019; Hotokezaka & Nakar 2020). One may define a thermalisation efficiency quantity $f(t)$ such that:

$$\dot{q}_{tot}(t) = \sum_i \dot{Q}_i(t) f_i(t) \text{ erg s}^{-1} \text{ g}^{-1}$$  \hspace{1cm} (3.3)

where $\dot{q}_{tot}(t)$ is the total power deposited at time $t$, $\dot{Q}_i(t)$ is the radioactive energy generation rate for a particular decay process, and $f_i(t)$ is the thermalisation efficiency of the decay product, i.e. what fraction of $\dot{Q}_i(t)$ is actually deposited to the KN. The thermalisation efficiency depends on the whole history of decays and degradations, as well as the geometry, density and composition of the ejecta (Barnes et al. 2016; Kasen & Barnes 2019; Waxman et al. 2019; Hotokezaka & Nakar 2020). This quantity allows decays from earlier times that only thermalise later to also be taken into account.

From Equation 3.3, two quantities need to be described for every decay product in...
order to calculate the ingoing energy to the KN. Though the exact form and evolution of these quantities remains under debate, a widely used $\beta$-decay energy generation rate for solar abundance ejecta at early times is (Metzger et al. 2010; Rosswog et al. 2018; Hotokezaka & Nakar 2020):

$$\dot{Q}_\beta(t) \approx 10^{10} t_d^{-1.3} \text{erg s}^{-1} \text{g}^{-1}$$

where $t_d$ is the time after merger in days. This energy generation rate includes all the possible products from $\beta$-decay: electrons/positrons, gamma rays, and neutrinos. Lower energy electrons from nuclear deexcitation, as well as $\sim$ keV X-rays may also be produced, but are found to have negligible contributions (Barnes et al. 2016). It should be noted however, that Equation 3.4 is sensitive to composition and nuclear network calculations (Rosswog et al. 2017; WanaJo 2018; Barnes et al. 2021; Zhu et al. 2021). Notably, the amount of power generated at 1 day after merger (i.e. $10^{10} \text{ erg s}^{-1} \text{g}^{-1}$) is found to range over an order of magnitude, as the unstable isotopes created by nuclear network calculations can vary significantly in species and quantity.

Alongside the raw energy generation rate, the thermalisation efficiency of the various $\beta$-decay products, i.e. electrons/positrons, gamma rays and neutrinos, must be calculated. Generally, neutrinos escape freely without depositing any energy, while gamma rays only thermalise efficiently at very early times (as shown in Figure 3.4 above). Electrons/positrons from the $\beta$-decay are thus expected to dominate the energy deposition from this decay channel.

Thermalisation efficiency is generally complex due to its evolution with time, and dependence on the geometry of the ejecta, as well as assumed properties of the thermalising decay product and ejecta composition. Extensive studies have been carried out on how non-thermal particles such as $\beta$-decay electrons/positrons lose their energy to the plasma in the ejecta (Huba 2013; Barnes et al. 2016; Waxman et al. 2019). Notably, these non-thermal electrons are expected to deposit their energy to the medium by plasma losses to free electrons (heating), electron impact ionisation, and atom excitation (see Barnes et al. 2016, and references therein). In KNe with media typically singly to doubly ionised, heating losses are expected to dominate (see e.g. Paper I). External factors such as the degree of magnetic field trapping of charged particles can also play a role in the heating efficiency of non-thermal electrons/positrons (Waxman et al. 2019; Hotokezaka & Nakar 2020).
Although the exact thermalisation efficiencies of non-thermal electrons/positrons must be calculated numerically, Kasen & Barnes (2019) provide a useful ad-hoc interpolation formula that links the high efficiency early times, to the later epochs where dropping densities reduce the thermalisation efficiency, given by:

\[ f_e(t) \approx (1 + t/t_e)^{-n} \]  

(3.5)

where the exact values of the thermalisation timescale \( t_e \) (note that “thermalisation timescale” is the timescale at which thermalisations become inefficient, not the time it takes to thermalise), and the power law index \( n \) vary depending on the assumed ejecta geometry, non-thermal electron emission energy, and degree of magnetic field trapping in the ejecta. Generally, these values lie in the approximate range of \( t_e \sim 15 - 40 \) days, and \( n \sim 1 - 1.5 \), for generic ejecta parameters e.g. \( M_{ej} = 0.05 \, M_\odot \), \( v_{ej} = 0.2 \, c \) (Barnes et al. 2016; Kasen & Barnes 2019; Hotokezaka & Nakar 2020). Thermalisation timescales in denser, slower moving ejecta such as those from disc winds are long, while those in faster moving dynamical ejecta can become very small, e.g. \( t_e \sim 5.6 \) days for \( M_{ej} = 0.01 \, M_\odot \), \( v_{ej} = 0.3 \, c \).

The wide range associated to \( t_e \) is due to various studies’ differing and definitions of thermalisation timescale (e.g. Hotokezaka & Nakar 2020). Qualitatively, \( t_e \) is defined as the timescale on which non-thermal electron thermalisation starts to become inefficient, or when the effective optical depth is unity (see e.g. Kasen & Barnes 2019; Hotokezaka & Nakar 2020, for definitions of thermalisation timescale). The variation of the power law index \( n \) depending on the degree of magnetic field trapping is explored in Paper I.

The gamma rays originating from \( \beta \)-decay are treated differently, and deposit energy by Compton scattering. They cease to thermalise efficiently when they can escape the ejecta without interactions, i.e. absorption or scattering, with a thermalisation efficiency evolving as (Barnes et al. 2016; Hotokezaka et al. 2016):

\[ f_\gamma(t) \approx 1 - e^{-(t_\gamma/t)^2} \]  

(3.6)

where \( t_\gamma \) is the thermalisation timescale for gamma rays.
The total energy deposited to the KN from $\beta$-decay can thus be written as:

$$\dot{q}_\beta = \dot{Q}_\beta (p_e f_e(t) + p_\gamma f_\gamma(t))$$ (3.7)

where $p_e \sim 0.2$ and $p_\gamma \sim 0.5$ are estimates of the fraction of $\beta$-decay energy going into non-thermal electrons/positrons and gamma rays respectively, thereby implying a fraction for neutrinos of $p_\nu \sim 0.3$ (Barnes et al. 2016).

### 3.3.2 $\alpha$-decay and Fission

The energy deposition associated to $\alpha$-decay is more heavily debated. Uncertainties in nuclear network quantities, such as fractions of $\alpha$-decay energy going into neutrinos, $\gamma$-rays and $\alpha$-particles vary from model to model (Barnes et al. 2021; Zhu et al. 2021). For analytical parametrisation of $\alpha$-decay deposition, there is also some debate as to whether the form should be that of an ensemble decay as for $\beta$-decay, or dominated by several single isotopes (see e.g. Kasen & Barnes 2019; Hotokezaka & Nakar 2020). Should the electron fraction be low enough, $Y_e \lesssim 0.2$, $\alpha$-decaying nuclei may be present in sufficient quantities to affect the energy deposition (Rosswog et al. 2017; Wu et al. 2019), with radioactive power $\dot{Q}_\alpha \sim 0.05 - 0.4 \dot{Q}_\beta$ (Kasen & Barnes 2019). Should many translead nuclei be present, the radioactive power law will follow that of $\beta$-decay in Equation 3.4, albeit with a possibly shallower index of $\dot{Q}_\alpha \propto t^{-1/d}$. However, if only a few isotopes are present, the $\alpha$-decay power will be a sum of several exponential decay laws. Some studies have shown that the $\alpha$-decay chains of only a few isotopes, notably of Rn ($Z=86, A \sim 222$) and Ra ($Z=88, A \sim 226$) may dominate $\alpha$-decay heating in KNe (see e.g. Wu et al. 2019). Should this be the case, then the power from these decay chains will be given by (Hotokezaka & Nakar 2020):

$$\dot{Q}_\alpha \approx 4 \times 10^8 \sum_i e^{-t/\tau_i} \left( \frac{Y_i}{10^{-5}} \right) \left( \frac{\tau_i}{10 \text{ days}} \right)^{-1} \left( \frac{E_i}{30 \text{ MeV}} \right) \text{ erg s}^{-1} \text{ g}^{-1}$$ (3.8)

where $Y_i$ is the initial number fraction of the parent nuclide per nucleon, $\tau_i$ is the mean isotope lifetime, and $E_i$ is the total energy release for that decay chain.

The thermalisation efficiency of $\alpha$-decay particles will also be different to that of non-thermal electrons. Kasen & Barnes (2019) use the same form as in Equation 3.5, though with a power law index of $n = 1.5$, and thermalisation timescale of $t_\alpha = 3 t_e = 45 - 120 \sim 37$.
days to account for higher plasma loss rates of $\alpha$-particles. This value is however highly sensitive to assumptions made about ejecta geometry, opacity, and the average energy of the $\alpha$-particles, as Hotokezaka et al. (2016) favour only the smaller value of $t_\alpha \approx 45$ days. This wide range of values highlights how critical the definition of thermalisation timescale is, and showcases the sensitivity of late time KN LCs to the contribution of $\alpha$-decays (Rosswog et al. 2017; Wu et al. 2019), which themselves are sensitive to the employed nuclear mass formula. $\alpha$-decay may also provide a non-negligible amount of helium in the ejecta (see Figure 3.2), which has been suggested as potentially significant for KN emission (Perego et al. 2022; Tarumi et al. 2023).

Fission may play an important role for the energy deposition of a KN in two ways: one directly by the thermalisation of fission products, and the other more indirectly by affecting the abundance of elements which undergo $\alpha$-decay. In either case, the creation of extremely heavy, neutron rich, and unstable nuclei is necessary, which requires a low electron fraction of $Y_e \lesssim 0.15$. Though the exact abundances of these elements remains largely unclear due to a lack of experimental data on abundance yields (Wanajo et al. 2014; Wanajo 2018; Vassh et al. 2019), spontaneous fission releases large amounts of energy, and thus can contribute significantly to the energy deposition. Should certain long lived isotopes such as $^{254}$Cf be present, they may even come to dominate late time radioactive heating (Wanajo et al. 2014; Wanajo 2018; Wu et al. 2019), with other isotopes such as $^{259}$Fm and $^{262}$Fm also potentially having important contributions at earlier times (Wanajo et al. 2014). Spontaneous fission will be dominated by a few isotopes, with the power contribution given by a similar formula to Equation 3.8, albeit with different normalisations (Hotokezaka & Nakar 2020).

While Papers I and II used simple, single-zone, 4-element solar abundance compositions, Paper III employed different composition models based on the considerations of ejecta properties described above. The electron fractions $Y_e \sim 0.35, 0.25, 0.15$ were chosen to represent a broad range of compositions and associated energy depositions, e.g. the $Y_e \sim 0.15$ model had significant $\alpha$-decay and fission contributions. Similarly, the total ejecta masses of $0.05 M_\odot$ and velocity range of $0.05 - 0.3 c$ were chosen to reproduce the bulk ejecta properties. Though these remain simplified models due to the usage of homogeneous compositions, they still represent the first step towards more ‘realistic’ inputs of ejecta model in nebular phase KN modelling, and remain the most complete models used in the NLTE regime thus far.
4 Electromagnetic Emission

After the merger completes, and the two compact objects coalesce, the electromagnetic emission begins. From theoretical considerations and modelling, several different types of emission are predicted. The earliest to occur is precursor emission from free neutron decay, which is theorised to be produced in the first minutes to hours. Alongside this early emission, a GRB is also expected to arise from the launching of a relativistic jet. As the jet reaches the ISM, shocks will occur, generating a non-thermal afterglow from synchrotron emission. Also on a timescale of the first few hours, unstable r-process isotopes will start to decay, powering the broad-band KN, which is the main object of this thesis. The merger ejecta will also eventually reach the ISM, at various times depending on the component velocity, which may also undergo shocks and thereby produce non-thermal KN afterglow emission. I will briefly describe here some further details of the precursor and KN afterglow emissions, which thus far have not been observed and therefore remain theoretical. The GRB and subsequent afterglow, as well as the KN are covered in more detail later in this Chapter, with a particular focus on the observables of AT2017gfo.

The first EM emission to occur is expected to be precursor emission from free neutron $\beta$-decay in the fast moving shocked dynamical ejecta that expand fast enough to prevent neutron capture (see e.g. Bauswein et al. 2013b; Metzger et al. 2015; Metzger 2019). The timescale of this process is short (minutes to hours as mentioned above), and may provide a significantly larger amount of early time heating than r-process radioactive decay on these timescales. This may substantially enhance early time emission if sufficient low density ejecta is present to allow free-neutron decay to be significant. However, such precursor emission has not yet been observed, and it is likely to be difficult to do so. Observation of such early time emission requires the detection and localisation of the merger event at its first moments, which may make observing precursor emission difficult even if robust theoretical models of the emission are established.
On much later timescales, and alongside the GRB afterglow, KN afterglows have been predicted to occur. These are theorised to be produced when the fastest moving dynamical ejecta in turn reaches the ISM (e.g. Nakar & Piran 2011; Hotokezaka & Piran 2015; Nedora et al. 2023). As these move through the ISM, shocks may be generated, which can produce non-thermal afterglow emission in a phenomenologically similar way to GRB afterglows, i.e. by synchrotron emission. Depending on the ejecta velocity and how relativistic it is, emission may occur on a timescale ranging from tens of days, to years (Hotokezaka et al. 2018; Nedora et al. 2021), from X-ray to radio wavelengths following synchrotron emission evolution. KN afterglows from fast moving dynamical ejecta may encode information about the early merger dynamics, notably pertaining to the nuclear equation of state for NSs, and thus may be interesting to observe. So far, however, no KN afterglow has been robustly detected, possibly due to the observational difficulties stemming from low luminosities and long timescales involved. However, theoretical work on making accurate KN afterglow emission models is ongoing, with a particular focus on the radio emission in the context of the Very Large Array (VLA) and upcoming Square Kilometer Array (SKA) (see e.g. Nedora et al. 2023, for a recent work).

4.1 Gamma Ray Bursts

Though little electromagnetic emission aside from precursor emission is expected before the actual merger event, any compact object merger including at least one NS may potentially launch a relativistic jet which can produce a GRB. The first observable EM emission is thus expected to be a GRB, typically of the 'short' variant lasting up to 2 seconds (see Nakar 2007; Berger 2014, for reviews), and generally occurring a few seconds after the merger. In the case of GW170817, the GRB 170817A was detected 1.7 seconds after the GW signal, in a sky area consistent with that marked out by the LIGO and Virgo interferometers (see Figure 2.5), and with a duration of 2.0 seconds as shown in Figure 4.1 (Abbott et al. 2017a; Goldstein et al. 2017; Savchenko et al. 2017). The detection of a GRB and its X-ray afterglow consistent with the GW merger signal allowed the event to be located on the sky, and thereby enabled complete EM follow up observations to be conducted on the BNS merger (Abbott et al. 2017a).

Short GRBs have been suggested to be produced by BNS and BHNS mergers for many years, with diverse supporting lines of evidence, both theoretical and observational (see
Figure 4.1: Binned lightcurve for GRB 170817A from the Fermi/GBM satellite, in the 50 - 300 keV band, where the red horizontal line is the background level. Note that T0 here is the time since the first detection of the GRB, and not the GW. Taken from Goldstein et al. (2017).

e.g. Eichler et al. 1989; Narayan et al. 1992; Katz & Canel 1996; Berger 2014; Fong et al. 2014), while long GRBs with lifetimes \( \gtrsim 2 \) seconds are associated to the core collapse of massive stars (see e.g. Kouveliotou et al. 1993; Nakar 2007; Jespersen et al. 2020), though this has recently been shown to not be consistently true (Gompertz et al. 2023; Gillanders et al. 2023b; Levan et al. 2023). In particular sGRBs are thought to require the presence of a massive accretion disc around the NS or BH remnant left after the merger. Since such a disc should form a few seconds after the merger (see Section 3.2.3), the launching of the sGRB associated to the merger is expected to follow right after the GW signal, which was indeed the case for GRB 170817A.

The exact details surrounding the origin of GRB 170817A have been theorised by many different groups (see e.g. Granot et al. 2017; Gottlieb et al. 2018; Beloborodov et al. 2020). All agree however, that the GRB was observed at a relatively large viewing angle \( \theta_{\text{obs}} \sim 23 \) deg (Finstad et al. 2018; Abbott et al. 2019), with a jet opening angle of approximately \( \theta_{\text{jet}} \sim 5 \) deg (Fong et al. 2017; Beniamini et al. 2019). Several explanations have been offered for to explain the observation of an off-angle GRB like this: a strong off-axis GRB along the rotation axis \( \sim 30 \) deg away (Kasliwal et al. 2017), a structured jet with a broad weak wing, or a cocoon shock breakout (Gottlieb et al. 2018). The physical processes related to the emission of sGRBs, and GRBs in general, remain an area of intense
research. Though they are known to be produced by the tightly collimated, relativistic jets emitted from the merger, the exact launching process is unclear. A commonly accepted theory is that the accretion disc is responsible for powering the jets (Ruffert & Janka 1999). Several popular mechanisms exist to explain where the jet and GRB energy comes from (see Gottlieb et al. 2023, for a review).

Though the processes associated to the launching of jets remain debated, it is clear that the sGRBs associated to NS mergers provide a useful tool for the localisation of the subsequent KN and broadband emission, even if observing the GRB itself can be difficult due to the anisotropic nature of the emission. For the majority of compact object mergers, the GRB emission will be beamed away from the line of sight, following that of the relativistic motion of the jet (Ryan et al. 2015; Metzger 2019). However, the jet eventually enters a causal regime as it slows down and shocks the ISM, thereby causing a delayed GRB afterglow across various wavelengths due to synchrotron emission (Troja et al. 2017; Margutti et al. 2017; Hallinan et al. 2017). Observing such afterglows allows information about the structure, orientation and energy of the jet to be obtained, which can in turn provide information about the physical state of the merger and remnant (Lamb & Kobayashi 2017; Nakar & Piran 2017; Lazzati et al. 2018).

In what has been a surprise to the community, recent works have linked long GRBs to compact object mergers. Notably, the 50-second long GRB 211211A at a distance of 350 Mpc (Rastinejad et al. 2022; Gompertz et al. 2023), and 35-second long GRB 230307A at a distance of 290 Mpc (Gillanders et al. 2023b; Levan et al. 2023) were much longer than 2 seconds and yet have been linked to NS mergers due to excess emission in their afterglows. This excess emission was found in the IR, with emission following that predicted by KN models at the observed epochs, implying that these GRBs were produced by compact object mergers. Levan et al. (2023) estimate that such bright GRB events with favourable beaming alignments occur with rates on the order of $R \sim 3 - 5 \times 10^{-3} \text{Gpc}^{-3}\text{yr}^{-1}$. No GW signals were associated to these objects however, due to the LIGO/Virgo observatories being offline at that time, making the presence of a compact object merger unverifiable by the medium of GWs. If the afterglow interpretation is correct, then the dichotomy of short and long GRBs associated to compact object mergers and stellar collapse respectively may not be as clear as once thought, hinting at the potential diversity of BNS mergers and their associated EM transients.
4.1.1 GRB Afterglows

The case of GRB 170817A provides an excellent, and thus far unique example of how GRB afterglows can be used to infer jet and merger properties. Initially at early times ($t \lesssim 2$ days), deep X-ray and radio non-detections allowed GRB 170817A to be set apart from other known GRBs. This was followed by a gradual rise of the lightcurve $F_\nu \propto t^{0.8}$ up to about $t \sim 150$ days with emission well modelled by a mildly relativistic cocoon outflow (Haggard et al. 2017; Hallinan et al. 2017; Ruan et al. 2018), with a sharp achromatic peak around that time. After this peak, a steep decline was seen, with the flux scaling as $F_\nu \propto t^{-2.2}$ at $t \gtrsim 300$ days, better modelled by emission from the jet core (Margutti & Chornock 2021). The evolution of the GRB afterglow lightcurves is shown in various bands in Figure 4.2.

For this event, an X-ray afterglow was initially detected by the Chandra space telescope approximately 9 days after the merger (Haggard et al. 2017; Margutti et al. 2017; Troja et al. 2017), with radio emission following later (Hallinan et al. 2017). Since the afterglow is produced by synchrotron emission, it is expected to emit at successively longer wavelengths as time goes on. This implies that an early time optical afterglow component exists too, though these wavelengths are expected to be dominated by the isotropic KN emission, thereby making the GRB optical afterglow difficult to observe before the KN component sufficiently fades (Lazzati et al. 2018; Lyman et al. 2018).

Many theories have been offered to describe GRB 170817A and its afterglow, with observational support from various wavebands, though some have now been eliminated by very late time observations. A jet and cocoon model, with a built-in mechanism for dissipating energy into gamma rays is capable of self-consistently explaining the observations (Bromberg et al. 2018; Gottlieb et al. 2018; Nakar & Piran 2018). Alternatively, dissipative processes in jet wings around an ultra-relativistic jet core may also account for observations (Lazzati et al. 2017, 2018; Mooley et al. 2018). Other models have however been eliminated from the late time behaviour of the GRB afterglow, such as quasi-spherical models with magnetar-like flares (Salafia et al. 2018), interaction of the fastest dynamical ejecta with the environment (Hotokezaka et al. 2018), and pure cocoon systems, which remained valid until the LC peak at 160 days when the jet core afterglow began to dominate emission (Kasliwal et al. 2017; Nakar & Piran 2018; Margutti & Chornock 2021). However, the exact details of the jet propagation, especially with regards to interaction with the...
merger ejecta, ISM, and winds emanating from an central engine or disc remain under intense study (see e.g. Lazzati et al. 2021; Nativi et al. 2022).

KNe have also been inferred from excess emission in the afterglows of GRBs. Rastinejad et al. (2021) lead a comprehensive study of sGRB afterglows with respect to possible KN excess emission, finding that afterglow observations were enough to constrain ejecta masses and velocities in 14 out of 85 objects. KNe were also inferred from the afterglow observations of the long GRBs 211211A and 230307A as mentioned above, with JWST taking spectra of the latter, at 29 and 61 days after the merger, shown in Figure 4.3 (Gillanders et al. 2023b; Levan et al. 2023). At these epochs, these KN is expected to be firmly in the nebular phase, which is the key topic of this thesis. By considering which forbidden transitions are strongest at these wavelengths and for the estimated densities at that time, suggestions for the two emission features were made. These were also proposed in the context of the 10 day spectra of AT2017gfo, and were the forbidden lines [Te III] 2.15 micron, and [W III]/[Se III] at 4.6 micron respectively. This represents the only (potential) KN observation made by JWST so far, and showcases a promising alternative method of detecting KNe that does not rely on a GW detection and localisation. This

Figure 4.2: GRB 170817A afterglow lightcurves in radio (pink and orange), optical (purple), and X-ray (blue), with model curves from various studies. Taken from Margutti & Chornock (2021).
Figure 4.3: JWST spectra of the GRB 230307A counterpart at 29 and 61 days, with a comparison to AT2017gfo at 10 days, and a 29 day model. The two emission features in the 29 day spectrum are tentatively identified as forbidden transitions in Te III, W III and/or Se III. Taken from Levan et al. (2023).

provides additional impetus for usage of scanning sky surveys to (proverbially) keep an eye out for long GRBs that do not fit the traditional characteristics of those associated to stellar collapse.

### 4.2 The Kilonova

Alongside the GRB and the associated afterglows, a NS merger is expected to produce a radioactively powered KN (or alternatively macronova) (Li & Paczyński 1998; Metzger et al. 2010; Rosswog et al. 2013). The various ejecta described in Section 3.2 are heated, ionised and excited by the deposition of radioactive decay energy, as described in Section 3.3. This is then translated into EM emission across the ultraviolet (UV), optical and infra-red (IR), with the exact spectral energy distribution (SED), magnitude and LC evolution depending on various factors. I begin by discussing a key quantity: opacity, which governs how radiation transports through the KN ejecta, both in terms of time and wavelength. Onwards from there, I describe how LC and SED modelling employ opacity to analyse KN emission. This section is of particular relevance to Paper II, where we investigate the importance of calculating such opacities with full NLTE and radiative
transfer calculations. I then take a closer look at how these observables can be analysed in order to derive properties of the merger.

4.2.1 Opacity

Two key quantities in astrophysics relating to any kind of radiative transfer are the opacity and optical depth. While opacity $\kappa$ is a local quantity, optical depth $\tau$ is applied more broadly to the entire KN. When the optical depth at certain wavelengths is too high ($\tau \gtrsim 1$), it is described as optically thick, and radiation of that wavelength cannot immediately escape the ejecta. Initially, the KN ejecta are dense, and optically thick over a wide range of wavelengths, from the UV to the NIR, so photons across these wavelengths diffuse out of the ejecta. This typically characterises the 'diffusion' phase of the KN, where most of the EM emission will come from the surface layer of the ejecta where $\tau \lesssim 1$. This surface layer can be approximated to emit like a blackbody (BB) photosphere, so this phase is also referred to as the photospheric phase. The peak EM emission from the KN can only be expected once the density drops enough such that photons can escape on a timescale similar to or smaller than the expansion timescale (more on this below).

Several processes can contribute to opacity, including free-free absorption from ionised gas, bound-free absorption (photo-ionisation), bound-bound (line) absorption, as well as electron scattering. The exact calculations of opacity and optical depth depends on the process responsible for it:

- **Free-free opacity**: arises from the absorption of photons by free electrons in the fields of positively charged ions, also known as inverse bremsstrahlung. Can be calculated exactly in a frequency dependent form if composition is known (see e.g. Equation 15.29 of Shu 1991)

- **Bound-free opacity**: arises from the photoionisation of atoms and ions by photons with energies greater than the ionisation energy of the level. Opacity calculation requires knowledge of photoionisation cross-sections, which are poorly known for r-process elements. Kramer’s bound-free opacity may be used as an approximation (see Carroll & Ostlie 2006, for a review in an astronomy context).

- **Bound-bound opacity**: arises from line transitions when a bound electron absorbs a photon with energy matching the transition energy. The electron is excited, and a
line absorption is produced. For KNe, the bound-bound opacity is often calculated from the Sobolev optical depths of the lines, appropriate for expanding media.

- **Electron scattering**: arises from photons scattering off free electrons. Can be important if the gas is highly ionised. At low photon energy, this is Thompson scattering. Klein-Nishina corrections are applied when photon energies are higher ($\geq m_e c^2$). Scattering off bound electrons can occur for very high photon energies with wavelengths smaller than the atomic scale.

In KNe, bound-bound opacity from extremely dense forests of line transitions in heavy r-process elements is found to dominate opacity at UV, optical, and near NIR wavelengths (e.g. Kasen et al. 2013; Tanaka & Hotokezaka 2013; Tanaka et al. 2020). In particular, lanthanide and actinide elements with an extremely high number of levels provide millions of line transitions contributing their optical depths to the overall ejecta opacity (Fontes et al. 2020; Tanaka et al. 2020; Silva et al. 2022; Flörs et al. 2023).

Since line opacity typically dominates for wavelengths relevant to KN emission, I focus now on the calculations for this process. Line opacity in an expanding medium with a velocity gradient, and thermal line widths much smaller than the expansion velocity, is calculated using the tried and tested Sobolev optical depth formalism (Sobolev 1960), which for homologous expansion ($R = v_{ej} t$) equals:

$$\tau^S = \frac{\pi e^2}{m_e c^2} f n_l \left( 1 - \frac{n_u g_l}{n_l g_u} \right) t \lambda_0$$  \hspace{1cm} (4.1)

Where $e$ is the electron charge, $m_e$ the mass of the electron, $c$ the speed of light, $f$ is the oscillator strength of the transition, $n_l$, $n_u$ are the lower and upper level number densities, with associated statistical weights $g_l$ and $g_u$ respectively, $\lambda_0$ is the line center rest wavelength, and $t$ is the time after merger.

The expansion opacity, which has been broadly applied in the context of SNe and KNe (Karp et al. 1977; Eastman & Pinto 1993; Kasen et al. 2006, 2013; Tanaka et al. 2020), is then calculated from a summation of the line optical depths:

$$\kappa_{\text{exp}}(\lambda, t) = \frac{1}{c \rho(t)} \sum_i \frac{\lambda_i}{\Delta \lambda(\lambda)} \left( 1 - e^{-\tau^S_i(t)} \right)$$  \hspace{1cm} (4.2)

Where the sum is conducted over all the lines within the wavelength bin $\Delta \lambda$. As seen from Equations 4.1 and 4.2, the calculation of Sobolev expansion opacity requires knowledge
of the excitation and ionisation structure of the ejecta, through \( n_u \) and \( n_l \). Naturally, the opacity is highly composition dependent, as the number of optically thick lines contributing to the opacity will vary immensely depending on the structure of the species present in the ejecta.

The details of this calculation, as well as its sensitivity to accurate level population and ionisation structure calculations in LTE vs NLTE is studied in Paper II. There, we also check the effect of different bin sizes \( \Delta \lambda \) on the opacity calculation, finding \( \Delta \lambda = 0.01 \lambda \) to be within the limits set by the ejecta velocity \( u_{ej}/c = 0.1 \) and thermal broadening of \( \Delta_{\text{therm}} \sim 10^{-6} \), and consistent with the values used in other studies (Kasen et al. 2013; Tanaka et al. 2020). Limitations of the Sobolev approximation and expansion opacity are discussed in a numerical context in Chapter 6.

In the past years, several compilations of expansion opacities have been calculated (e.g. Tanaka et al. 2020; Carvajal Gallego et al. 2023; Banerjee et al. 2023). These studies conducted a thorough and systematic calculation of LTE expansion opacities for the vast majority of r-process elements. They find that the average ejecta opacity is heavily composition dependent, which can be roughly parametrised by electron fraction \( Y_e \), as described in Section 3.1. The range of opacity values is significant, from \( \kappa_{\text{exp}} \approx 20 - 30 \text{ cm}^2 \text{ g}^{-1} \) for the lowest \( Y_e \lesssim 0.2 \), down to \( \kappa_{\text{exp}} \approx 1 \text{ cm}^2 \text{ g}^{-1} \) for \( Y_e \approx 0.4 \). Depending on the wavelength and epoch, lanthanide and actinide expansion opacities for a given element may reach values as high as \( 10^3 \text{ cm}^2 \text{ g}^{-1} \) in the blue optical \( \sim 3000 \text{ Å} \), while still reaching \( 10 \text{ cm}^2 \text{ g}^{-1} \) at 2.5 micron (see e.g. Fontes et al. 2023; Flörs et al. 2023).

These expansion opacities are high, with only the most lanthanide poor value coming close to Fe rich Type Ia SNe, which have maximal values of around \( \kappa \approx 1 \text{ cm}^2 \text{ g}^{-1} \) (Pinto & Eastman 2000). The exact values of the calculated opacity depend on certain assumptions made, as well as the method used to calculate them. Notably, existing expansion opacities all assume LTE conditions for ionisation and excitation in the ejecta, which may not be a good approximation past a few days after merger. This assumption is tested in Paper II, with the consequences with respect to LC and SED modelling, and ejecta parameters inferred from those such as lanthanide fraction, explored therein. Other approaches to ejecta opacity include a ‘line-smeared’ approach (Fontes et al. 2017; Wollaeger et al. 2018), which find even higher lanthanide opacities than with the expansion opacity formalism (see Fontes et al. 2020, for a comparison of opacity methods). Many studies have previously also used a ‘gray’ opacity, which is averaged over all wavelengths
and broadly applied to the ejecta (see e.g. Drout et al. 2017; Metzger 2019; Collins et al. 2023). I discuss the limitations of using ‘gray’ opacity further below.

4.2.2 Lightcurves and SEDs

The lightcurve of the KN traces how much light is emitted over time, and can be considered either as a total integral across all wavelengths (bolometric LC), or in individual wavelength bands. The LCs by photometric band for AT2017gfo are shown in Figure 4.4 as an example of a complete LC set from the near UV to the NIR.

The key to understanding how these LCs evolve, when they rise, peak and fall, lies with the evolution of the ejecta opacity and the optical depth. Consider a very simple ejecta model that is expanding spherically and homologously. The overall optical depth can be estimated simply by using a constant, wavelength averaged ‘gray’ opacity (Metzger 2019):

$$
\tau \approx \rho \kappa_{\text{exp}} R = \frac{3 M_{\text{ej}} \kappa_{\text{exp}}}{4 \pi R^2} \approx 70 \left( \frac{M_{\text{ej}}}{0.01 M_\odot} \right) \left( \frac{\kappa_{\text{exp}}}{1 \text{ cm}^2 \text{ g}^{-1}} \right) \left( \frac{v_{\text{ej}}}{0.1c} \right)^{-2} \left( \frac{t}{1 \text{ day}} \right)^{-2}
$$

(4.3)
Right after the initial explosion, the ejecta are extremely hot, but the thermal radiation cannot escape as the optical depth across all wavelengths is too high ($\tau \gg 1$), even for relatively 'low' opacity lanthanide-free ejecta. This corresponds to a long photon diffusion time through the medium ($t_{\text{diff}} \gg t$):

$$t_{\text{diff}}(t) \approx \frac{R\tau}{c} = \frac{3M_{\text{ej}}\kappa_{\text{exp}}}{4\pi cv_{\text{ej}}t}$$  \hspace{1cm} (4.4)

When the diffusion time is greater than the evolutionary time, the KN is said to be in the 'diffusion' phase. Since densities are high and the ejecta conditions are most likely close to LTE, there is in effect a photosphere emitting as a blackbody at a temperature of $T_{\text{phot}}$. Photons emitted there escape into line forming regions further out into the ejecta, where they interact with species, typically by scattering processes, and form P-Cygni lines.

Ejecta temperatures for the observed KN AT2017gfo were inferred from the observed SEDs and their colours by assuming a BB spectrum (see Figure 4.5) (e.g. Drout et al. 2017; Arcavi 2018; Waxman et al. 2018). However, this method requires the ejecta to be firmly in LTE conditions, which may not be true after several days (see Paper II). Similarly, reprocessing of the radiation field by fluorescence and resonance scattering may engender a temperature difference between the gas and photons (see Shingles et al. 2023, and paper III).

In order for radiation to be able to effectively escape the ejecta, the diffusion timescale must drop until it is comparable to the evolutionary time, i.e. $t_{\text{diff}} \sim t$. This condition defines the characteristic time at which the bolometric LC will peak (Arnett 1982). Thus, setting $t_{\text{diff}} = t = t_{\text{peak}}$ in the above equation, and solving yields:

$$t_{\text{peak}} = \left( \frac{3M_{\text{ej}}\kappa_{\text{exp}}}{4\pi cv_{\text{ej}}} \right)^{1/2} \approx 1.6 \text{ days} \left( \frac{M_{\text{ej}}}{0.01 M_\odot} \right)^{1/2} \left( \frac{\kappa_{\text{exp}}}{1 \text{ cm}^2 \text{ g}^{-1}} \right)^{1/2} \left( \frac{v_{\text{ej}}}{0.1c} \right)^{-1/2}$$  \hspace{1cm} (4.5)

The bolometric luminosity at the LC peak can be estimated using Arnett’s Law (Arnett 1982):

$$L_{\text{peak}} \approx M_{\text{ej}}\dot{q}(t_{\text{peak}})$$  \hspace{1cm} (4.6)

Where $\dot{q}$ is the total heating rate including thermalisation efficiency, as described in Section 3.3. As such, the calculations of peak luminosity are sensitive to the radioactive decay
Figure 4.5: Example of blackbody fits to photometric data of AT2017gfo. Taken from Waxman et al. (2018).
power as calculated by nuclear networks, as well as the plasma physics describing the effective thermalisation of the various decay products. Using Equation 4.6, and assuming ejecta mass, velocity and a 'gray' opacity, the peak luminosity is found to be (Metzger 2019; Cowan et al. 2021; Margutti & Chornock 2021):

$$L_{\text{peak}} \approx 10^{41} \text{ erg s}^{-1} f(t) \left( \frac{M_{\text{ej}}}{0.01 \text{ M}_\odot} \right)^{1-\alpha/2} \left( \frac{v_{\text{ej}}}{0.1c} \right)^{\alpha/2} \left( \frac{\kappa_{\text{exp}}}{1 \text{ cm}^2 \text{ g}^{-1}} \right)^{-\alpha/2}$$ (4.7)

where $f(t)$ is the thermalisation efficiency (see Section 3.3), expected to be near unity at peak times, but which may rapidly drop for certain decay products (e.g. gamma-rays), and $\alpha \sim 1.3$ stems from the ensemble radioactive $\beta$-decay of r-process isotopes (see Section 3.3).

An important caveat to the simple analysis presented here lies with the usage of a constant, wavelength averaged 'gray' opacity. In reality, the opacity varies significantly with both time and wavelength (see e.g. Tanaka et al. 2020, and Paper II). From Figure 4.4, we also see that the LCs of AT2017gfo evolved differently depending on waveband. Since the ejecta opacity are composition dependent, different compositions will yield different peak times, which can be used to infer ejecta compositions. Typically, ejecta with heavier compositions arising from lower values of $Y_e$ have higher opacity, such that ejecta with relatively lanthanide free, or high $Y_e$ compositions will have LCs that peak on the order of a couple of days, whereas lanthanide rich ejecta will have LCs peaking much later, on the scale of a week (assuming the range of opacities from Tanaka et al. (2020)).

The significance of the wavelength dependency is compounded by the high probabilities of fluorescence following each line absorption. Notably, opacity tends to increase with time as densities drop (see Paper II for an example), implying that calculations of a peak LC time from equations such as Equation 4.5 may be poorly defined for KNe. However, observations of AT2017gfo show clear peak times and luminosities, and so combined with numerical calculations of opacity from theoretical atomic data, can help constrain the values of opacity at early times.

**Inferred properties of AT2017gfo**

As was previously shown in Figure 4.4, the observed LCs of AT2017gfo varied significantly in amplitude and evolution in different wavebands. The UV wavebands peaked first, and already exhibited a declining luminosity at the first observation around $t = 0.65$ days.
The various optical wavebands peaked around 1 day, with redder wavebands having slightly later peak times, and NIR wavebands such as the $K$-band peaking around 3.5 days after merger. The different peak times and amplitudes across various bands imply that the KN ejecta may have been split into various components with different compositions (see e.g. Cowperthwaite et al. 2017; Kasen et al. 2017; Perego et al. 2017).

From the consideration of various hydrodynamic simulations, it is suggested that two to three ejecta components are required in order to fully account for the observed emission: dynamical ejecta, secular ejecta, and wind outflows from the disc (e.g. Perego et al. 2017). Taking all the observations into account, component masses, velocities and opacities were estimated to be (Villar et al. 2017): a blue component with $M_{ej} \approx 0.020 \, M_\odot$, $v_{ej} \approx 0.27 \, c$, $\kappa_{\text{exp}} = 0.5 \, \text{cm}^2 \, \text{g}^{-1}$, an intermediate opacity "purple" component with $M_{ej} \approx 0.047 \, M_\odot$, $v_{ej} \approx 0.15 \, c$, $\kappa_{\text{exp}} = 3 \, \text{cm}^2 \, \text{g}^{-1}$, and a red component with $M_{ej} \approx 0.011 \, M_\odot$, $v_{ej} \approx 0.14 \, c$, $\kappa_{\text{exp}} = 10 \, \text{cm}^2 \, \text{g}^{-1}$.

The bluest component is typically referred to as the lanthanide-free component, and is expected to have a very small lanthanide fraction $X_{\text{lanth}} < 10^{-4}$ (Kasen et al. 2017), corresponding to a high electron fraction $Y_e \gtrsim 0.3$ (e.g. Rosswog et al. 2018) (see Figure 3.2). Contrastingly, the red, lanthanide-rich component would have much more significant lanthanide fractions, estimated on the order of $X_{\text{lanth}} \sim 10^{-3} - 10^{-2}$ (Kasen et al. 2017; Tanaka et al. 2017; Waxman et al. 2018), with the purple component thus having a lanthanide fraction somewhere in between.

The initial blue component that peaked at very early times has been associated to the fastest ejecta layers in the polar direction, containing only light r-process elements (Drout et al. 2017; Nicholl et al. 2017). The SED transition to redder colours at later times suggested lanthanide-rich ejecta originating from slower moving disc ejecta, consistent with larger masses and slower velocities (Siegel & Metzger 2017; Kasen et al. 2017; Perego et al. 2017; Fernández et al. 2019). Though the collapse time of the AT2017gfo remnant was initially poorly known, more recent hydrodynamical simulations have estimated the collapse time to be on the order of $t \gtrsim 0.1 \, \text{s}$ (Just et al. 2023; Kawaguchi et al. 2023), based on models broadly reproducing the observed LCs. Combining the contributions from all various ejecta, and sources estimating their properties, a total ejecta mass in the range of $M_{ej} \sim 0.03 - 0.08 \, M_\odot$ was found (e.g. Kasen et al. 2017; Cowperthwaite et al. 2017; Perego et al. 2017; Villar et al. 2017).
Around the peak time and moving forwards, the ejecta will become optically thin at
more and more wavelengths, due to decreasing densities as the medium expands. Eventu-
ally, when the diffusion time is small relative to the evolutionary time for all wavelengths,
the KN will enter the steady-state ‘nebular’ phase. At this point, the bolometric lightcurve
will track the instantaneous energy deposition, and the spectrum transitions to be more
dominated by emission lines. For AT2017gfo, the only very late time observations of the
KN were photometric measurements by the *Spitzer Space Telescope*, at 43 and 74 days
after merger, seen in the 4.5 \( \mu \text{m} \) band, but not in the 3.6 \( \mu \text{m} \) band (Villar et al. 2018;
Kasliwal et al. 2022). Such observations of KNe are interesting, as they probe radioactive
heating at those epochs, which may provide constraints on thermalisation of decay prod-
ucts, as well as the type of decay product at these late times. This in turn can provide
hints as to which elements and isotopes are present, particularly in the case where only
a few isotopes are theorised to dominate energy deposition (e.g. \(^{223}\text{Ra}, ^{225}\text{Ra}, ^{225}\text{Ac}, ^{254}\text{Cf}\)
Wanajo et al. 2014; Wanajo 2018; Wu et al. 2019). Inferring this requires understanding
the physical conditions of these ejecta at late times, as well as the details of the physics
in the nebular phase (see Chapter 5).

**4.2.3 Spectral Analyses of AT2017gfo**

Aside from analysing the LCs and broad SED shape of AT2017gfo, much effort has been
conducted in analysing the spectra, taken by the X-Shooter instrument and Very Large
Telescope (VLT) (Pian et al. 2017; Smartt et al. 2017), and available from 1.5 to 10.4
days after merger. Since most of these epochs are within the photospheric or diffusion
phase, they are expected to resemble BB continua with P-Cygni lines overlain. Therefore,
the analysis of many early time spectral studies has focused on the various P-Cygni
features found in the spectra. Some of the first spectral analyses conducted on AT2017gfo
made use of the radiative transfer code **TARDIS** (Kerzendorf & Sim 2014, see Chapter 7
for description of code), which, making use of limited atomic data from Kurucz (2018),
tentatively attributed some early spectral features to neutral tellurium (Te I) and caesium
(Cs I) (Smartt et al. 2017). The modelling was conducted in LTE, using a BB continuum
fit with some absorption lines implemented from the atomic data, yielding preliminary
results that provided some first insights into the potential of element identification offered
by spectral analysis, as shown in Figure 4.6. Later works ruled out these species on the
Figure 4.6: AT2017gfo spectra from 1.4 to 4.4 days, with photometric data (left panel). On the right panels, TARDIS model spectra are compared to the observed spectra, with important lines from Cs I and Te I marked out. These broadly reproduce key absorption features seen in early time spectra, providing a first ever spectral analysis of a KN. Taken from Smartt et al. (2017).

basis that the absorption lines of these species were not strong enough to produce such features at the estimated ejecta temperatures (Watson et al. 2019), while others found that singly ionised species tended to dominate in the ionisations structure (Domoto et al. 2021; Gillanders et al. 2022).

The analysis of the spectra of AT2017gfo continued in many subsequent studies once the spectral data set was complete up to 10.4 days. These studies typically analysed the spectra with considerations of LTE physics, such that fitting BB continua to determine temperature evolution was a common approach (see e.g. Pian et al. 2017; Waxman et al. 2018). Combined with the photometric evolution of the transient, showing a notable ‘blue to red’ colour evolution, ‘2-component’ KN models as described above were generated to explain the evolution of AT2017gfo. These typically consisted of a fast moving, low mass, blue or lanthanide-free component, and a slower moving, heavier, red or lanthanide-rich component (Kasen et al. 2017; Tanaka et al. 2017; Villar et al. 2017; Waxman et al. 2018).
A fit using a single component model is shown in grey, while the AT2017gfo spectrum is shown as a solid black line. The values in bracket show the lanthanide mass fractions of the models. Taken from Waxman et al. (2018), which used the models of Kasen et al. (2017).

2018; Wollaeger et al. 2018; Metzger 2019). The emission from the two components was typically added together to form a composite spectrum. This allowed the LC evolution of AT2017gfo, and key spectral features, notably the 2.1 micron ‘bump’ emerging at later times as shown in Figure 4.7, to be explained.

Subsequently to these initial works, the first broadly accepted element identification was found by Watson et al. (2019). In this paper, several approaches, both analytical and using RT codes such as TARDIS, were used to study a P-Cygni feature with rest wavelength of $\sim 1$ micron. This was shown to be well reproduced by singly ionised strontium (Sr II), which has a powerful triplet transition at $\lambda_0 \sim 1$ micron. This triplet had the proper characteristics to explain the P-Cygni feature visible up to $\sim 5$ days after merger, as shown in Figure 4.8. Since then, Sr II has been widely accepted as present in the ejecta of AT2017gfo, with further studies supporting this finding (e.g. Domoto et al. 2021, 2022; Gillanders et al. 2022, 2023a).

Though Sr II is one of the more broadly accepted species identifications in AT2017gfo, it is not unanimously agreed upon that it is solely responsible for the 1 micron P-Cygni feature. An alternative, or complimentary explanation considering neutral helium (He
Figure 4.8: Identification of Sr II in the ejecta of AT2017gfo using a P-Cygni line analysis. Taken from Watson et al. (2019).
I) emission was also suggested (Perego et al. 2022; Tarumi et al. 2023). These studies found that treating the ionisation and excitation state of He I requires inclusion of NLTE processes, notably non-thermal (NT) collisional ionisation. The excited level of He I potentially contributing to the \(~1\) micron emission has an energy that is beyond the reach of thermal collisional excitation for estimated ejecta temperatures of \(~5000\) K, and is instead populated by recombination. Since the ionisation potential of He I is also too high to be thermally ionised, NT collisional ionisation is required to have a sufficiently high abundance of He II needed to provide adequate recombination flow to the excited He I state. There, it was found that the evolution of the P-Cygni feature is consistent with a contribution from both Sr II and He I, with Sr II dominating at earlier times, and He I taking over at later times (in the range of 1.4 - 4.4 days). However, this result is also dependent on the amount of photoionising UV flux redistributed by heavier species, which was only approximately modelled in Tarumi et al. (2023). A deeper analysis requires a full fluorescence treatment to properly model, such as that used in Paper III.

The P-Cygni style analysis was also used by several other studies, with the most similar direct identification of another element by Sneppen & Watson (2023). In this analysis, it was suggested that the faint P-Cygni signature visible at \(~760\) nm from 3 - 5 days may be due to several singly ionised yttrium (Y II) lines. An alternative explanation from neutral rubidium (Rb I) is suggested in Paper III, with an explanation as to why therein.

Other studies have considered more broadly the impact of various species and their expansion line opacities on the overall shape of the spectrum (see e.g. Gillanders et al. 2021; Domoto et al. 2021, 2022; Gillanders et al. 2023a). A couple works have employed Bayesian inference to recover abundance patterns from the spectra of AT2017gfo, recovering the Sr II P-Cygni feature, as well as making predictions for other first r-process peak and lanthanide species (Vieira et al. 2023a,b). More complex modelling of the early time spectra using various 3D RT codes has also been conducted by several groups, where it is often found that viewing angle plays a large role in the emergent spectrum (e.g. Bulla 2019; Wollaeger et al. 2021; Collins et al. 2023), implying that spherically symmetric, 1D models may not be capturing important geometrical effects.

At later epochs of available data, when the spectra are noisier and NLTE effects begin to play a role, identification of elements may be less straightforward, and typically is more dependent on the accurate treatment of various NLTE processes (see Chapter 5 for a physical overview of these processes). The atomic data pertaining to these processes
is typically poorly known, if at all, forcing the usage of approximate treatments such as fitting formulae derived from the values measured for lighter elements. This makes the reliable and robust identification of spectral features more complex and challenging.

However, some progress has also been made there, by determining which species may be strongly emitting from a consideration of atomic structure and ejecta gas state conditions. Modelling of NLTE effects in these later time studies varies, from full consideration of NLTE physics in dedicated RT codes, such as in Paper III of this thesis and Hotokezaka et al. (2021), to more analytical considerations of NLTE effects making use of various simplifying assumptions (Hotokezaka et al. 2022; Gillanders et al. 2023a; Hotokezaka et al. 2023). In Gillanders et al. (2023a), the evolution of the Sr II 1 micron feature was studied in detail up to 10 days after merger, assuming optical thinness. Using collisional data for Sr II transitions, as well as fixed temperatures estimated from earlier TARDIS models (Gillanders et al. 2022), and estimates of electron density from NLTE SN works (Jerkstrand 2017), the level populations of the lowest lying states relevant for the strong transitions were estimated. From these, line luminosities and profiles were calculated using Sobolev escape probabilities and experimentally measured transition strengths (A-values). This work also suggested potential candidates (La III, Ce III, Gd III, Ra II) for late time features following a simpler procedure considering the location of experimentally known strong lines, with an allowed E1 transition from Ce III being suggested as responsible for the 2.1 micron bump in the 10.4 day spectrum of AT2017gfo.

In Hotokezaka et al. (2022), potential candidate species responsible for the 43 and 74 day 4.5 micron band emission of AT2017gfo photometrically observed by Spitzer (Villar et al. 2018; Kasliwal et al. 2022) were examined. This work considered only forbidden magnetic dipole (M1) transitions, with fixed collision strengths of Υ = 1 (see Chapter 7 for a discussion), and fixed ionisation fractions for the species of key interest. These were chosen based on estimated M1 line strengths in the 2.5 - 6.0 micron wavelength range, relevant to the Spitzer observations. The ejecta were also assumed to be optically thin in this wavelength regime, which is likely to be the case for such late epochs. In doing so, it was found that forbidden doubly ionised selenium [Se III] and tungsten [W III] lines were both suitable candidates for this emission, depending on whether 2nd r-process peak elements were present or not in the ejecta.

A similar study was conducted later, focussing this time on the 10 day spectra of AT2017gfo and the 2.1 micron emission of a forbidden [Te III] line (Hotokezaka et al.
Figure 4.9: Potential explanation of the later spectra of AT2017gfo considering the forbidden lines of Te I and Te III. Other 2nd peak elements are included for reference, but are found to be weak in comparison. Taken from Hotokezaka et al. (2023).
In this study, the collision strengths were not fixed to unity, but instead taken from detailed atomic (r-matrix, Berrington et al. 1987) calculations following observations of the [Te III] line in planetary nebulae (Madonna et al. 2018). The [Te III] 2.1 micron line was thus suggested as a potential candidate for this emission bump observed in AT2017gfo as shown in Figure 4.9, offering an alternative explanation to that of Ce III proposed in Gillanders et al. (2023a). The forbidden [Te III], [Se III] and [W III] lines were also suggested in the context of the afterglow excess emission of the long GRB 230307A (see Figure 4.3 and Levan et al. 2023).

While these studies may be simpler in their approach than usage of a fully NLTE capable RT code such as SUMO, they perfectly exemplify what can be gained from such approaches. For these studies, the results of Paper I and II were directly considered in their analyses, showing that the testing of commonly used assumptions in the field provides a useful contribution to the KN modelling community. Paper III of this PhD thesis did not seek to identify individual features of AT2017gfo, but rather provide a broad overview of the spectral formation processes and SED shape. However, a strong dipole transition in Rb I was suggested as an alternative candidate for the 760 nm P-Cygni line suggested by Sneppen & Watson (2023), given the strong scattering behaviour observed across all epochs in our models. Though this candidate transition is far from robustly verified, this result highlights the utility of employing different methods to probe all the possible candidates for KN spectral formation.
5 The Nebular Phase

As the NS merger ejecta continue to expand, densities and optical depths continue to drop. Decreasing densities lead the thermodynamic conditions away from LTE. The ejecta will also eventually become optically thin at most wavelengths, at which point the KN is considered to be in the nebular phase. Several atomic lines may remain optically thick until later times, providing any remaining optical depth in the nebular phase (see Section 4.2.1 and Equations 4.1, 4.2). This also implies that the spectra will become emission line dominated, as opposed to having a quasi-blackbody continuum as in the photospheric phase (see Section 4.2.2). Furthermore, as the photon diffusion time (Equation 4.4) becomes short relative to the evolutionary time, the bolometric LC will track the instantaneous energy deposition as described previously in Section 3.3.

For KNe, the transition of conditions away from LTE to NLTE is expected to occur quite rapidly, and was estimated to be only a few days after the merger for AT2017gfo (e.g. Smartt et al. 2017), with spectral shapes becoming inconsistent with a BB spectrum around 7 days (see e.g. Waxman et al. 2018). However, as found in Paper III, the ejecta are not yet necessarily optically thin at this point. This thesis therefore provides the first detailed investigation as to when the transition to the optically thin nebular phase occurs, finding that is likely to be later than 20 days.

The nebular phase is interesting to study for various reasons. Firstly, as the spectra are expected to be emission line dominated, they provide an opportunity to identify (groups of) elements and their abundances in the ejecta. This can help constrain r-process yields from compact object mergers. As the LC in the nebular phase initially tracks instantaneous energy deposition, it can be used to gain more information on r-process radioactive decay chains, which may likewise help to constrain abundances of unstable isotopes. This may be particularly effective for long-lived species that may dominate late time energy deposition (e.g. Wanajo 2018; Wu & Banerjee 2022). Since
emission in this phase probes the entire ejecta, it also provides a chance to determine not only total element masses, but also their morphological distributions, thus tying in to the hydrodynamical models as overviewed in Section 3.2.

However, extracting such information from the nebular phase of KNe is far from trivial for several reasons. Alongside the drop in optical depth, decreasing ejecta density precludes LTE conditions, such that finding the nebular phase gas state requires NLTE solvers (e.g. Jerkstrand et al. 2011; Hotokezaka et al. 2021). At the start of the nebular phase, many lines remain optically thick, so radiative transfer is still needed, at least up to 20 days as shown in Paper III. Furthermore, much atomic data related to r-process elements, and necessary for realistic NLTE models of the nebular phase, are crucially lacking from experimental databases. This forces the use of atomic data codes (see e.g. Kasen et al. 2013; Fontes et al. 2017; Wollaeger et al. 2018; Kasen et al. 2017; Fontes et al. 2020; Tanaka et al. 2020, and Paper I) which rely on a host of assumptions regarding atomic properties in order to generate the requisite quantities, and as such possess inherent uncertainties. In many cases, cross-sectional data for the various NLTE processes at play in the nebular phase are lacking, both experimentally and theoretically, such that fitting formulae derived from the properties of lighter elements are employed instead.

Regardless of such challenges, modelling and understanding emission in the nebular phase remains of great interest, with promising rewards should it be accomplished. As such, understanding what key processes dominate, and how they can be implemented into simulations, is central to this endeavour. In this section, I describe the key physical attributes and processes related to the nebular phase of KNe. I consider the equations relevant to nebular phase modelling of explosive transients (based on Jerkstrand 2011, for a SN context), paying close attention to when the steady-state and time-dependent conditions are met, which were closely investigated in Paper I.

5.1 Physics of Nebular Phase Kilonovae

As the KN first enters the nebular phase, the atomic processes occurring within the ejecta are all fast relative to the evolutionary time, and thus the radioactive energy deposited is almost instantaneously re-emitted as light. These key processes determine the thermodynamic conditions of the ejecta: temperature, ionisation and excitation. In the following section, I will cover the equations relevant to these three key quantities, and the condi-
tions required on their associated timescales in order for the steady-state regime to apply, assuming that the energy deposition comes solely from radioactive decay, i.e. that no central engine or fall-back accretion is present.

5.1.1 Temperature

The temperature in the KN is a result of various heating and cooling processes working against each other, and can be calculated from a form of the first law of thermodynamics:

\[
\frac{dU(t)}{dt} = Q(t) - P(t) \frac{dV(t)}{dt}
\]

(5.1)

where \(U(t)\) is the internal energy of a segment of gas, \(P(t)\) the pressure, \(V(t)\) the volume, and \(Q(t) = H(t) - C(t)\) the net heating rate, given by the difference between the heating and cooling rates of diverse processes described below. Several steps can be taken in order to rewrite the above equation in a more convenient form:

- Assume an ideal mono-atomic gas: \(U(t) = \frac{3}{2} N(t) k_B T(t)\), where \(N(t)\) is the total number of particles.

- \(N(t) = N_a (1 + x_e(t))\), where \(N_a\) is the number of atoms/ions, and \(x_e(t)\) is the ionisation fraction.

- The ionisation fraction for a single species is given by \(x_e(t) = x^+(t) + 2x^{2+}(t) + 3x^{3+}(t)\), where the superscripts correspond to ionisation state. The total ionisation fraction is a summation over all species according to their respective mass fractions in the ejecta composition.

- \(P(t) = N(t) k_B T(t)/V(t) = n(t) k_B T(t)\), where \(n(t)\) is the total number density of particles, following the assumption of an ideal mono-atomic gas.

- Assume homologous expansion: volume is given by \(V(t) = V_0 (t/t_0)^{-3}\), yielding \(dV(t)/dt = 3V(t)/t\).

- \(Q(t)/V(t) = (H(t) - C(t))/V(t) = h(t) - c(t)\), where \(h(t)\) and \(c(t)\) are heating and cooling rates per volume, respectively.
Going from Equation 5.1, the first law of thermodynamics can then be written as:

\[
\frac{dT(t)}{dt} \approx \frac{h(t) - c(t)}{3/2k_B n(t)} - \frac{2T(t)}{t} - \frac{T(t)}{1 + x_e(t)} \frac{dx_e(t)}{dt}
\]  

(5.2)

From the above equation, we see that the change in temperature is not only dependent on the heating and cooling rates at the current epoch, but also directly on the current temperature solution, as well as the ionisation fraction and its evolution. The last term we call the "ionisation cooling" term, though is typically expected to be small compared to the others (see details in the appendix of Paper I). The \(2T/t\) term is the adiabatic cooling term, which is expected to play a role later in the time-dependent phase (see Section 5.1.4), and is investigated in Paper I.

In the steady-state nebular phase, the thermal equilibrium approximation is applied, where \(h(t) = c(t)\), and temperature is solved solely from this algebraic equation, as opposed to the full IVP (Initial Value Problem) equation. If one assumes that heating is purely from radioactivity, with non-thermal particles counting as radiation, and that cooling is entirely dominated by radiative processes (as opposed to conductive processes), this approximation also corresponds to radiative equilibrium in the co-moving frame. Naturally, solving \(h(t) = c(t)\) represents omitting the rest of the terms in Equation 5.2, so it is important to understand when this approximation can be made.

The key lies with the timescales associated to the heating and cooling processes relative to the timescale describing how quickly physical conditions, such as density or temperature change. Assuming that the "ionisation cooling" term in Equation 5.2 is negligible compared to the other terms (shown in Paper I), we can rearrange the equation to isolate the net heating term, with the steady-state condition then being written as:

\[
\frac{dT(t)}{dt} + \frac{2T(t)}{t} < < \begin{cases} 
h(t)/(\frac{3}{2}n(t)k_B) \\
(\frac{3}{2}n(t)k_B)c(t)/(t) 
\end{cases}
\]

(5.3)

In order to express the above condition in terms of timescales, we can divide everything by temperature and take the reciprocal, such that the condition becomes:

\[
\left(\frac{T(t)}{T(t)} + \frac{2}{t}\right)^{-1} > > \begin{cases} 
(\frac{3}{2}n(t)kB_T(t))/h(t) \equiv t_{\text{heat}} \\
(\frac{3}{2}n(t)kB_T(t))/c(t) \equiv t_{\text{cool}} 
\end{cases}
\]

(5.4)

where the term on the right-hand side is either the cooling or heating timescale, and
the term on the left-hand side is the change timescale. The temperature can only change significantly on times equal to or longer than the radioactive decay timescale or the expansion timescale. We choose to compare the expansion timescale to the cooling timescale, as this is a more direct comparison of the relative importance of adiabatic and line cooling, an aspect central to Paper I. For this timescale, the approximation \( \dot{T}(t) \sim T(t)/t \) is used, which yields a timescale for change of \( t_{\text{change}} = t/3 \). Steady-state conditions for temperature will hold if the cooling/heating timescales are (significantly) shorter than this. However, the accuracy of this approximation is not always very good, as shown in Figure 5.1, which implies that the comparison between the change and heating/cooling timescales may sometimes be very approximate. In any case, in order to verify that steady-state holds for the temperature equation, we must first understand which heating and cooling processes dominate in the ejecta.

**Heating Processes**

Heating may generally occur from a variety of processes including non-thermal particle collisions from radioactive decay, photoionisations, collisional de-excitations after line absorptions, and free-free absorptions. In the nebular phase, heating is expected to be dominated by energy deposition from radioactive decay in the form of non-thermal collisions (e.g. Li & Paczyński 1998), as is the case in nebular phase SNe (Kozma & Fransson 1992; Jerkstrand 2011). From the models run with SUMO, we find non-thermal plasma heating to entirely dominate. As mentioned previously, the radioactive energy deposition is split into heating, ionisation and excitation by solving the Spencer-Fano equation (Spencer & Fano 1954; Kozma & Fransson 1992), and we find that heating fractions are typically between 95 – 99% in our KN models.

If we ignore the small difference of this factor to unity, the total heating rate per volume can thus be estimated by considering the input from the energy deposition described in Section 3.3, divided by the expanding volume of the ejecta:

\[
h(t) = M_{ej} \dot{q}_{\text{tot}}(t)/(4\pi v_{ej}^3 t^3) \text{ erg s}^{-1} \text{ cm}^{-3}
\]

where \( \dot{q}_{\text{tot}} \) is the total radioactive energy per mass deposited to the ejecta, as written in Equation 3.3, \( M_{ej} \) and \( v_{ej} \) are the ejecta mass and velocity respectively, and we assume homologous expansion.
Figure 5.1: Checking the approximation $d(T)/dt \sim T(t)/t$ in the derivation of the change timescale for steady-state temperature, for our Paper I model with $M_{\text{ej}} = 0.05M_\odot$ and $v_{\text{ej}} = 0.1c$, run in steady-state mode. Early epochs show that this approximation may be off by up to an order of magnitude, showcasing that the definition of the change timescale as $t_{\text{change}} \sim t/3$ is not always accurate. In this case $dT(t)/dt$ is always positive, however this does not generally have to be the case. Should this quantity be negative, the change timescale will be $t_{\text{change}} < t/2$, which highlights a further potential issue with this approximation. Reused from Licentiate Thesis.
Cooling Processes

Cooling likewise occurs from various processes, such as line cooling, recombination, and free-free emission. As for SNe, cooling for nebular phase KNe is expected to be line dominated, especially due to the presence of line rich r-process elements (Tanaka et al. 2020; Hotokezaka et al. 2021). Indeed, looking at the cooling contributions of these processes in our SUMO models, we find that line cooling typically makes up $\gtrsim 99\%$ of the contribution, with recombination second on the order of $\lesssim 1\%$. As such, in the following analytical calculations, we assume that the $c(t)$ term in Equation 5.2 is completely line cooling dominated. Adiabatic cooling at later times may become important as dropping densities mitigate line cooling, though this term is independently included in the energy equation 5.2, and is not part of the cooling term $c(t)$. The contribution of adiabatic expansion to cooling, and its effect on the bolometric LC are studied in Paper I.

Generally, the cooling from line emission is not directly equivalent to the total emission in that line, as the bound electron excitation preceding the radiative decay must come from a collision with a thermal electron. As such, one must also consider how thermal electrons lose and gain energy by collisional (de)-excitation. These details are examined thoroughly in Paper I, but I still write here a simple form of the line cooling for an ion $i$ to demonstrate its intricate nature:

$$c^i_{\text{line}}(t) = \Lambda^i(T, n_e, n_i)n_e n_i \text{ erg s}^{-1} \text{ cm}^{-3}$$

(5.6)

where $\Lambda^i(T, n_e, n_i)$ describes the line cooling capacities of the ion, $n_e$ is the electron number density and $n_i$ is the number density of that ion. Generally, this function also depends on the radiation field by photoionisation flows, as well as the number density of the next ionisation stage $n_i^{+}$ by recombination flows, and time by the line optical depths, thereby greatly increasing its complexity. The total line cooling for the ejecta is found by summing over ion fractions, i.e. $\Lambda(T, n_e, n_1, ..., n_l) = \sum_i x_i \Lambda^i(T, n_e, n_i)$, where $x_i$ is the number fraction of the ion, given by the ion fraction multiplied by the element’s number fraction in the model.

At low enough densities, the line cooling function is expected to become solely a function of temperature, and increases monotonically provided all other quantities are constant. This can be used for a simplistic estimation of the evolution of temperature for nebular phase KNe, which is found to increase with time post-peak (see Hotokezaka et al.

At low enough densities, the line cooling function is expected to become solely a function of temperature, and increases monotonically provided all other quantities are constant. This can be used for a simplistic estimation of the evolution of temperature for nebular phase KNe, which is found to increase with time post-peak (see Hotokezaka et al.
Given these intricacies, it is difficult to map out line cooling functions in a systematic manner, though we attempt to do so in the low density regime for the ions included in our models in Papers I and II. We generally find that ion species with large level and line counts, such as the lanthanides and actinides, are stronger coolers than their lighter r-process counterparts. This follows the trend that complex species such as lanthanides also provide the most opacity to the ejecta (see Section 4.2.1).

**Conditions for Steady-State Temperature**

With the cooling and heating terms now well defined, we can try to establish a distinct condition for the application of the thermal equilibrium approximation, where temperature is found by solving \( h(t) = c(t) \) in Equation 5.2. Before plugging the heating and cooling expressions from Equations 5.5 and 5.6 into the heating and cooling timescales of Equation 5.4, it is convenient to write out the expressions for ion and electron number density, for a uniform sphere model:

\[
\begin{align*}
n_i(t) &= \frac{M_i}{A_i m_p} \left( \frac{4\pi}{3} v_{ej}^3 t^3 \right)^{-1} \\
n_e(t) &= \frac{M_{ej} < A > m_p}{A > m_p} \left( \frac{4\pi}{3} v_{ej}^3 t^3 \right)^{-1} x_e(t)
\end{align*}
\]

where \( M_i \) is the mass of the element in the model, \( M_{ej} \) is the total ejecta mass, \( A_i \) is the atomic mass number of the element, \( < A > \) the average atomic mass number of the ejecta composition, and \( m_p \) is the mass of a proton. The total number density is then the sum of all ion densities and the electron density.

With these written out, we can now plug the expression for heating per volume into the heating timescale (Equation 5.4), and simplify in order to find an expression for \( t_{\text{heat}} \):

\[
t_{\text{heat}} = \frac{3}{2} k_B T(t) \left( \frac{1 + x_e(t)}{< A > m_p q_{\text{tot}}(t)} \right)
\]

In the NLTE regime for KNe, we expect the temperature and ionisation fraction \( x_e \) to initially increase, while energy deposition \( q_{\text{tot}} \) naturally decreases due to decreasing power generation and thermalisation efficiencies. Thus, the heating timescale increases with time at a rate faster than \( q_{\text{tot}}(t)^{-1} \propto t^{1.3} \), and will eventually become comparable
to the expansion time estimated by $t/3$. The same approach can be used for the cooling
time-scale, yielding:

$$t_{\text{cool}} = \frac{3}{2} k_B \frac{T(t)}{\Lambda(T, n_e, n_1, ..., n_l) n(t)} \frac{1 + x_e(t)}{x_e(t)}$$ (5.10)

The evolution of the cooling timescale is more difficult to estimate, as various terms
evolve in opposite ways. From our models, we find that temperature in the steady-state
nebular regime tends to increase with time, and the number density $n(t)$ will decrease
with time, thus increasing the cooling timescale. The ionisation fraction $x_e(t)$ tends to
increase slowly, so the fraction $(1 + x_e(t))/x_e(t)$ will slightly decrease the cooling timescale.
However, the main unknown is the line cooling function, which increases with temperature
but decreases with density. Past a certain critical density however, it is expected to become
solely a function of temperature, and thus will likely increase, following temperature.
Thus, we find that the cooling timescale will also eventually become comparable to the
expansion time $\sim t/3$. This may be at very late times for certain models however, as
shown in Figure 5.2.

The exact critical time(s) when the heating/cooling timescales become long relative
to the expansion time will depend on how various quantities such as temperature, ion-
isation and heating/cooling evolve, and thus is difficult to estimate analytically. What
is unambiguous however, is that the steady-state approximation for temperature, where
$h(t) = c(t)$ is used to solve for the ejecta temperature, requires such heating and cooling
timescales to be (much) shorter relative to the expansion time. Paper I has a more in
depth analysis of when the cooling timescale may become long, and thus $h(t) \neq c(t)$,
implying that the steady-state approximation can no longer be used.
Figure 5.2: Comparison of the heating and cooling timescales to the change timescale $t_{\text{change}} = t/3$ for our standard model $M_{eJ} = 0.05 \, M_\odot$, $v_{eJ} = 0.1c$, run in time-dependent. For this model, the heating and cooling timescales remain much smaller than the change timescale for all epochs considered here, and are almost identical to each other ($t_{\text{cool}}$ becomes very slightly smaller than $t_{\text{heat}}$ at late times). As such, this model is expected to have temperature well described by the steady-state calculation, and the approximation $h(t) = c(t)$ is therefore well applied. Reused from Licentiate Thesis.
5.1.2 Ionisation

The ionisation structure of the KN ejecta is determined by ionisation and recombination processes working against each other. Since the ejecta in the KN are rapidly expanding and conditions move away from LTE, the Saha equation can in general not be used (see Paper II for details), and ionisation structure must instead be calculated from rate equations. Consider an ion \( i \) with number fraction \( x_i \), in ejecta with ionisation rate per particle \( \Gamma_i \), and recombination rate per particle \( \Psi_i \). The rate of change of this ion’s abundance is given by:

\[
\frac{dx_i}{dt} = \Gamma_{i-1} x_{i-1} - (\Psi_i + \Gamma_i) x_i + \Psi_{i+1} x_{i+1}
\]

This equation can be applied to every ion from every element, though naturally, neutral atoms will not have recombination to less ionised states. Due to restrictions in available data, we only allow up to triple ionisation, so the equation for the triply ionised ions will not have any further ionisation. In general, we find that our ionisation solutions do not reach this ‘roof’, and so expect this limitation to have negligible effects. The steady-state nebular phase allows the use of the ionisation equilibrium approximation, in which the time derivative on the left hand side is set to zero. Ionisation structure is then calculated from the following algebraic equations:

\[
\Gamma_{i-1} x_{i-1} = \Psi_i x_i
\]

\[
\sum_{i=0}^{3} x_i = 1
\]

\[
n_e = \sum_j n_j \sum_{i=0}^{3} i x_i
\]

where \( i \) is the index of the ion, \( j \) is the index for a particular element, \( Z(j) = 3 \) the maximal charge of that element (e.g. number of electrons lost), and \( n_j \) the number density of that element in the ejecta composition. The second equation exists for every element, and is a number conservation equation, while the third equation conserves charge.
The steady-state approximation is only valid if the time derivative, or change of ion fraction for a particular ion is small relative to the other terms in Equation 5.11:

\[
\left| \frac{dx_i}{dt} \right| \ll \begin{cases} 
\Gamma_{i-1} x_{i-1} + \Psi_{i+1} x_{i+1} \\
(\Psi_i + \Gamma_i) x_i 
\end{cases}
\]  

(5.15)

As for temperature, the above conditions will be fulfilled either together or not at all, provided the ionisation and recombination timescales are small compared to the timescale for significant ionisation structure change. Using the second expression, we find:

\[
\left( \frac{|dx_i(t)/dt|}{x_i(t)} \right)^{-1} \gg \frac{1}{\Gamma_i + \Psi_i}
\]  

(5.16)

As for temperature, we compare the change timescale to the evolutionary time by assuming \(|dx_i(t)/dt| \sim x_i(t)/t\) (see Figure 5.3). There, we see that the approximation is typically poor at early times, with significant deviations even at later times, motivating the detailed investigation of steady-state applicability carried out in Paper I. The ionisation and recombination timescales are simply defined as the reciprocals of the total ionisation and recombination rates i.e. \(t_{\text{ion}}^i = 1/\Gamma_i\), \(t_{\text{rec}}^i = 1/\Psi_i\). From these definitions, it is clear that the ionisation and recombination timescales will be equal to or longer than \(1/(\Gamma_i + \Psi_i)\), and so we may compare these directly to the change timescale \(t_{\text{change}} = t\) in order to verify that ionisation equilibrium is satisfied. This requires understanding of which processes dominate ionisation and recombination in KN ejecta.

**Ionisation Processes**

Ionisation in **SUMO** for KN models occurs by non-thermal (NT) particle collision (from radioactivity), thermal collisional ionisation (TI), and photoionisation (PI). Physically, the NT collisions may occur for any of the outer shell electrons, as well as electrons from inner shells. Should ionisation from an inner shell occur, subsequent processes may be induced, such as the Auger process, and X-ray fluorescence. As such, the total contribution of non-thermal collisional ionisation to the overall ionisation rate should take these factors into account. However, cross sections for inner r-process electron shells are lacking, and the Auger effect is currently not implemented in **SUMO**. Therefore, the non-thermal ionisation rate is calculated using the ionisation potential of the outermost valence electron. We
Figure 5.3: Checking the approximation $dx_i(t)/dt \sim x_i(t)/t$ using Te II as a reference ion. Te II is chosen as it is the most abundant ion for much of the epochs in our standard Paper I model with $M_{ej} = 0.05 \ M_\odot$, $v_{ej} = 0.1c$. The approximation appears poor at early times, but improves at later times. Reused from Licentiate Thesis.
also assume that his leaves the ion in the ground state after collision. The non-thermal ionisation rate is therefore given by:

$$\Gamma_{\text{NT}}^i = \frac{\dot{E}_i}{n_i^{\text{gm}} \chi_i^{\text{valence}}} = \frac{\rho_e(t) \dot{q}_{\text{tot}} S F_{\text{ion},i}(t)}{n_i^{\text{gm}}(t) \chi_i^{\text{valence}}}$$  \hspace{1cm} (5.17)

where $\dot{E}_i$ is the energy per second, per volume going into the ionisation of ion $i$, $n_i^{\text{gm}}(t)$ is the number density of the ground-multiplet (currently we do not include non-thermal ionisation from excited states), $\rho_e$ is the ejecta density, and $\chi_i^{\text{valence}}$ is the ionisation potential of the valence electron. The total energy going into ionisation is a combination of the overall radioactive energy deposition $\dot{q}_{\text{tot}}$ per unit mass as described in Section 3.3, and $SF_{\text{ion},i}(t)$ is the fraction of non-thermal energy leading to ionisation for that ion, found from solving the Spencer-Fano equation, typically on the order of $1-5\%$ for all ionisations together.

The ionisation fraction also relies on an assumption for the non-thermal collisional cross sections. As for many aspects of r-process data, these values are mostly lacking for r-process elements and their ions. In an attempt to approximate reasonable cross sections for this process, we employ the formalism described in Lotz (1967):

$$\sigma(E) = \psi a \frac{1}{\chi^2} \ln \left( \frac{E}{\chi} \right) \frac{E}{\chi}$$  \hspace{1cm} (5.18)

where $\psi a$ is a slowly varying function of the electron energy $E$ in eV, where $a = 4 \times 10^{-14}$ and $\psi$ is set to unity, $\chi$ is the ionisation potential in eV. This typically yields a cross section on the order of $10^{-17}$ cm$^2$ for electron energies of 1 keV and an ionisation potential of 10 eV. Further details are described in Paper I.

Thermal collisional ionisation was not included in Papers I and II, but was implemented for Paper III. This is treated by using the formalism of Shull & van Steenberg (1982):

$$\Gamma_{\text{TI}} = 1.3 \times 10^{-8} F \xi \chi_{eV}^{-2} T^{1/2} \left( 1 + a \frac{kT}{\chi} \right) e^{-\chi/kT} \text{s}^{-1}$$  \hspace{1cm} (5.19)

where $\chi$ is the ionisation potential ($\chi_{eV}$ in electron-volts), $F$ is a focussing factor of order unity, $\xi$ is the number of equivalent electrons in the valence shell, taken to be unity, $a = 0.1$. While it was initially predicted in Paper II that TI played a minimal role, this was confirmed in Paper III, where the TI rates were found to be consistently several orders of magnitude smaller than the NT and PI rates at every epoch, and for every model.
Photoionisation may play a role if the radiation field is intense enough, and the ionisation threshold low enough compared to average photon energies, which is typically the case for neutral r-process ions, as well as singly ionised lanthanides and actinides. The photoionisation rate of a particular ion \(i\) is formally:

\[
\Gamma_i^{\text{PI}} = \sum_l y_l \int_{\nu_0(l)}^{\infty} \sigma_{i,l}(\nu) \frac{4\pi J_\nu}{h\nu} d\nu
\]  

(5.20)

where \(y_l\) is the fraction of ion \(i\) in excitation state \(l\), \(\sigma_{i,l}(\nu)\) is the photoionisation cross section, \(\nu_0(l)\) is the threshold frequency for ionisation, and \(J_\nu\) is the mean intensity of the radiation field. Due to the lack of photoionisation cross sections for r-process elements, we currently use hydrogenic cross sections (Rybicki & Lightman 1979) in order to calculate photoionisation rates. For a level \(l\) in ion \(i\), the PI cross section for a photon with frequency \(\nu\) is given by:

\[
\sigma_{i,l}^{\text{PI}} = 7.9 \times 10^{-18} \left( \frac{\nu}{\nu_{i,l}^0} \right) g_{i,l} n_{i,l} Z_{i}^{-2}
\]  

(5.21)

where \(\nu_{i,l}^0\) is the ionisation threshold of the level, \(g_{i,l}\) the Gaunt factor of the level (set to unity for simplicity), \(Z_i\) the effective nuclear charge of the species, given by number of protons minus number of screening electrons. \(n_{i,l}\) is the effective principal quantum number given by:

\[
n_{i,l} = \left( 1 - \frac{E_{i,l}}{\chi_{i,0}} \right)
\]  

(5.22)

where \(E_{i,l}\) is the level energy, and \(\chi_{i,0}\) is the ground state ionisation potential. The PI rates in SUMO are estimated from the MC packet absorption in the radiative transfer step of the simulation, as described in Chapter 6 (see also Appendix of Jerkstrand et al. 2012).

Recombination Processes

Recombination can occur by collisional, radiative, dielectronic, stimulated recombination and charge transfer reactions. Typically, collisional recombination will be unimportant for post-peak SNe and KNe, being a 3-body reaction requiring high density. Charge transfer rates are unknown for r-process elements, but are generally less important for higher ionisation conditions. The key processes for KNe are expected to be radiative and dielectronic recombination, with stimulated recombination incorporated into photoionisation.
Radiative recombination occurs when an electron is captured by an ion and loses its kinetic energy by emitting a photon (this provides recombination cooling as mentioned in Section 5.1.1). The radiative recombination rate is related to a radiative recombination coefficient $\alpha^{rr}(T)$, and electron number density $n_e$ such that:

$$\Psi^{rr}_i = \alpha^{rr}_i(T)n_e$$ (5.23)

The radiative recombination typically scales as $\alpha^{rr} \propto T^{-1/2}$ for temperatures relevant to KNe ($T \gtrsim 5000$ K) (see e.g. Nahar et al. 1997).

Dielectronic recombination occurs when a captured electron transfers its excess energy to another bound electron, as opposed to radiating it away. This leaves the ion with two excited electrons, and does not emit radiation. This state is however not stable, and an electron may be ejected again in a process known as auto-ionisation, or may de-excite by emitting radiation. This process is typically important at high temperatures for most light elements, but has been found to potentially dominate recombination for r-process elements even at relatively low temperatures ($T \gtrsim 10^3$ K) (Hotokezaka et al. 2021). Similarly to radiative recombination, the dielectronic recombination rate can be related to an effective dielectronic recombination coefficient (omitting auto-ionisation fraction), and electron number density:

$$\Psi^{dr}_i = \alpha^{dr}_i(T)n_e$$ (5.24)

In both cases, however, these recombination coefficients are typically unknown for r-process elements and their ions. It is therefore difficult to adequately estimate recombination rates for the species present in KN ejecta. Obtaining values for these coefficients must therefore be done from numerical calculations by atomic codes, or by estimation from known values for lighter elements. The latter approach is used and explained in detail in Paper I, where we opt to take a fixed total recombination coefficient of $\alpha^{tot} = \alpha^{rr} + \alpha^{dr} = 10^{-11}$ cm$^3$ s$^{-1}$.
Condition for Steady-State Ionisation

Now that the ionisation and recombination processes are defined, the rates can be put into the condition for steady state in Equation 5.16. Let us first consider the recombination timescale, which after scaling for typical KN ejecta parameters, is found to be:

\[ t_{\text{rec}} \sim 2 \times 10^{-4} \, \text{days} \left( \frac{M_{\text{ej}}}{0.05 \, M_\odot} \right)^{-1} \left( \frac{v_{\text{ej}}}{0.1 \, c} \right)^3 \left( \frac{\alpha}{10^{-11} \, \text{cm}^3 \, \text{s}^{-1}} \right)^{-1} x_e(t)^{-1} t_d^3 \]  \hspace{1cm} (5.25)

where \( t_d \) is the time in days after the merger. A similar approach can be conducted for the non-thermal collisional ionisation, noting here that this time-scale is somewhat of an overestimate for the total ionisation time-scale as it omits photoionisation:

\[ t_{\text{NT},\text{ion}} \sim 2 \times 10^{-2} \, \text{days} \left( \frac{\text{SF}_{\text{ion}}}{0.01} \right)^{-1} \left( \frac{\bar{X}}{20 \, \text{eV}} \right) f(t)^{-1} t_d^{1.3} \]  \hspace{1cm} (5.26)

where \( f(t) \) is the overall thermalisation efficiency of the radioactive decay products (see Equations 3.3, 3.5, and 3.6), and we assume an average ionisation potential of r-process elements and ions of 20 eV.

Both timescales increase with time at a rate faster than linear, and thus will eventually become equal to and surpass the evolutionary time-scale, \( t \). Notably, the recombination timescale rises on the order of \( t^3 \) due to dropping densities and assuming \( x_e(t) \) to be roughly constant, and so will rapidly become long, especially for lower density dynamical ejecta. The ionisation timescale depends on the thermalisation efficiency of the radioactive energy deposition, which at early times is close to unity. However, at later times when thermalisation is inefficient, we have \( f(t) \propto t^{-1.5} \) (for \( \beta \) and \( \alpha \)-decay, see Paper I), and thus the non-thermal ionisation timescale rises on the order of \( t^{2.8} \), similar to the recombination timescale (see Figure 5.4).

A critical time when this occurs can be calculated by equating the ionisation or recombination timescales to the evolutionary time, which is done in detail in Paper I. Naturally, these critical times are sensitive to certain parameters, especially the recombination timescale which depends on our fiducial choice of recombination coefficient. The ionisation timescale is more robust in this sense, but the omission of photoionisation from the estimate also implies its value is somewhat overestimated. As such, these critical times are taken to be rough approximations. Similarly to temperature, however, the assurance that steady-state is applicable to ionisation structure is only valid if \( t_{\text{rec}}, t_{\text{ion}} << t \), which is
Figure 5.4: Comparison of the ionisation and recombination timescales for Te II in our one-zone toy model run in time-dependent mode with $M_{\text{ej}} = 0.05 \, M_\odot$, $v_{\text{ej}} = 0.1c$, to the change timescale $t_{\text{change}} = t$. The ionisation and recombination timescale are not systematically much smaller than the change timescale across all epochs considered here. Furthermore, the recombination timescale grows faster than the ionisation timescale. This separation can lead to time-dependent effects in the ionisation structure. Reused from Licentiate Thesis.
certainly likely to be the case the first few days after the merger. It will however become violated on time-scales of weeks or months, much faster than the time-scale of several years for SNe (Fransson & Kozma 1993).

5.1.3 Excitation

Similarly to ionisation structure, the excitation structure of a particular species within the ejecta is described by rate equations detailing processes into and out of the excitation state. Consider the fraction of an element in ionisation state $i$ and excitation state $l$:

$$\frac{dx_{i,l}}{dt} = \sum_{l'} x_{i-1,l'} \Gamma_{i-1,l',i,l} + \sum_{l'} x_{i+1,l'} \Psi_{i+1,l',i,l} + \sum_{l' \neq l} x_{i,l} \xi_{i,l,l'}$$

where $\Gamma$ denotes ionisation, $\Psi$ recombination, and $\xi$ internal transition rates respectively, and $l'$ is any excitation state $l \neq l'$. The internal transition rates include collisional and radiative excitations, whilst the ionisation and recombination rates are described in Section 5.1.2 above. As with ionisation, the excitation equilibrium approximation used in steady-state modelling corresponds to setting the time derivative on the left hand side to zero. As before, this approximation is only valid if the de-excitation timescale (or excitation time-scale, equivalently) is short compared to the evolutionary time. It is useful to look at the various internal transition rates to gauge under which circumstances this condition is satisfied.

Let us first split the internal transition rates into upwards and downwards transitions respectively, with upper level $u$ and lower level $l$:

$$\xi_{l,u} = R_{l,u}^{\text{abs}} + C_{l,u}^{\text{thermal}} + C_{l,u}^{\text{NT}}$$

$$\xi_{u,l} = R_{u,l}^{\text{spont}} + R_{u,l}^{\text{stim}} + C_{u,l}^{\text{thermal}} + C_{u,l}^{\text{NT}}$$

where radiative rates are given by $R$ and collisional rates are given by $C$. 81
Radiative Transition Rates

Radiative bound-bound transitions include absorption, spontaneous emission and stimulated emission. Starting with spontaneous emission, we must first consider the particular nature of the environment. In an expanding medium with differential velocity gradients, an extra escape mechanism for a photon emitted by spontaneous radiative de-excitation is available as the gradients shift atoms out of resonance with the newly emitted photon. However, other atomic lines may be shifted into resonance at different points in the medium by the same effect. The Sobolev approximation allows for this tricky situation to be solved when the expansion velocities are much larger than the intrinsic thermal line width (Sobolev 1960). In this approximation, the transfer through a given line becomes a purely local process, and the spontaneous emission rate is given by:

\[ R_{u,l}^{\text{spont}} = A_{u,l} \beta_{u,l} \]  
(5.30)

where \( A_{u,l} \) is the Einstein A-coefficient and \( \beta_{u,l} \) is the Sobolev escape probability. This quantity describes the angle and frequency averaged probability that a photon emitted by a line will escape the resonance layer, rather than be reabsorbed in the same line by another atom. It is defined in terms of the Sobolev optical depth, which for homologuous expansion is given by:

\[ \beta_{u,l} = 1 - e^{-\tau_{u,l}^S} \]  
(5.31)

where \( \tau_{u,l}^S \) is the Sobolev optical depth, the same quantity as described in Equation 4.1, and used in calculations of expansion opacity (see Section 4.2.1).

The radiative absorption rate in the Sobolev approximation is formally defined as:

\[ R_{l,u}^{\text{abs}} = B_{l,u} J_{\nu}^b \beta_{u,l}^S \]  
(5.32)

where \( B_{l,u} \) is the Einstein B-coefficient, and \( J_{\nu}^b \) is the radiation field in the blue wing of the line, where no line photons will be present. However, in SUMO these rates are estimated by counting Monte Carlo events, and \( J_{\nu}^b \) is not explicitly used. The stimulated emission rate is related to the absorption rate by the statistical weights of the levels:

\[ R_{u,l}^{\text{stim}} = \frac{g_l}{g_u} R_{l,u}^{\text{abs}} \]  
(5.33)
Collisional Transition Rates

Collision rates include thermal and non-thermal collisions, where thermal collisions can both excite and de-excite bound electrons. Non-thermal electrons may also de-excite, but this process is typically unimportant. Let us consider first the thermal collisional de-excitation rate (Osterbrock & Ferland 2006):

\[ C_{u,l} = 8.63 \times 10^{-6} \frac{\Upsilon_{u,l}(T_e)}{g_u T_{e}^{1/2}} \text{ cm}^3 \text{s}^{-1} \]  

(5.34)

which is directly related to the thermal collisional excitation rate by the statistical weights of the levels and a Boltzmann factor:

\[ C_{l,u} = \frac{g_u}{g_l} \zeta_{u,l} e^{-E_{u,l}/k_B T_e} \]  

(5.35)

In the above equations, \( T_e \) is the electron temperature, \( g_u \) and \( g_l \) are the statistical weights of the levels, and \( \Upsilon_{u,l} = \Upsilon_{l,u} \) is the dimensionless effective collision strength. The details of how \( \Upsilon \) is calculated in SUMO are described in Paper 1, and is generally different depending on whether the transition between the levels is allowed or forbidden (see van Regemorter 1962; Axelrod 1980; Rutten 2003; Botyánszki & Kasen 2017; Shingles et al. 2020).

Non-thermal collision rates also require knowledge of the collision strengths and interaction cross sections between high energy electrons and the ion species in the ejecta. Unfortunately, as for many r-process elements and data, these quantities are not known experimentally, and have yet to be thoroughly calculated numerically. When cross sections are unknown one may use the Bethe approximation, though this has thus far not been implemented for the r-process elements in SUMO. However, since typically 95–99% of the non-thermal energy goes to heating in KN conditions, and most of this is re-emitted by thermal collisional line excitations, additional direct non-thermal excitations are not expected to play an important role. This corresponds to having a non-thermal excitation energy fraction from the Spencer-Fano routine set to zero.

Conditions for Excitation Steady-State

Like for temperature and ionisation, the conditions required for steady-state to be applied to the excitation structure depends on the associated (de)excitation timescales relative to the evolutionary time. Considering the deexcitation rates, both collisional and radiative, it is extremely unlikely that the shortest timescale, e.g. for spontaneous radiative decay, will
be longer than $\sim 10^3$ seconds, which will always be shorter than the evolutionary timescale even at 1 day after the merger. Thus, it is expected that the excitation structure can always be calculated in steady-state, and thus the excitation equilibrium approximation is always applied. Paper I has a more detailed explanation of this argument in Appendix A.

5.1.4 The Time-Dependent Nebular Phase

I have covered in this Chapter the conditions in the nebular phase which are necessary in order for the steady-state approximation to be valid. In that regime, fast reprocessing timescales allow the various approximations to be applied to calculations of temperature, ionisation and excitation. However, as previously discussed, many of these timescales increase at rates faster than $t$, and thus will eventually become comparable or longer to the evolutionary time. At the point where the steady-state condition for these timescales, notably $t_{\text{scale}} \ll t$, is no longer satisfied, the nebular phase begins to transition into the so-called "time-dependent" phase.

Though the excitation structure calculations may always remain in the steady-state approximation due to relatively fast spontaneous decay rates, the temperature and ionisation structure must now be calculated using the full forms of Equations 5.2 and 5.11 respectively. For temperature, this means that processes such as adiabatic cooling, and "ionisation" cooling from the ionisation fraction dependent term now play a role, while ionisation structure must now account for the extra terms previously ignored. These extra terms can lead to a deviation of the bolometric LC from the instantaneous energy deposition, as reprocessing timescales become long, and energy may be lost to other processes such as adiabatic cooling.

In general, time-dependent effects can lead to different solutions than those of the steady-state calculations, and so it is interesting to determine when these deviations may begin to occur, and how significant they are. The time at which time-dependent effects become significant for temperature and ionisation structure can be estimated from critical times found by equating key timescales such as heating, cooling, recombination and ionisation timescales to the evolutionary time, i.e. $t_{\text{scale}} \sim t$. However, it is possible that these effects begin to have an effect before this point, as long as steady-state conditions are not rigorously satisfied. This subject is studied and described extensively in Paper I.
6 Kilonova Simulations with SUMO

In this chapter, I will describe in technical terms how SUMO conducts the spectral simulations of KNe. An overview of the radiative transfer (RT) methodology is also given, followed by a description of various numerical tests conducted to ensure the stability and validity of SUMO results.

6.1 Gas State Solutions

SUMO works to solve temperature, ionisation structure, and excitation structure self-consistently in the NLTE regime. These solutions are coupled to the radiation field by usage of a 1D Monte Carlo (MC) radiative transfer calculation (Jerkstrand 2011; Jerkstrand et al. 2011, 2012). Concretely, this is done by iterating through three main routines, corresponding to the statistical and thermal equilibrium calculations (suse), the radiative transfer (sumo), and the ionisation structure calculation (suib). A graphical representation of the overall flow of SUMO can be seen in Figure 6.1.

The suse and suib routines work by solving the various NLTE equations described in Section 5.1 for temperature, excitation, and ionisation. Temperature, excitation structure and ionisation structure are all found by solving the Equations 5.2, 5.11, and 5.27 defined in Chapter 5, iteratively using the Newton-Raphson method (Press et al. 1992). Temperature in the steady-state approximation is found by balancing the cooling and heating per volume terms (i.e. \( h(t) \) and \( c(t) \) in Equation 5.2), while time-dependent temperature is found using the same Newton-Raphson approach, but with the extra time-dependent terms and previous solutions now taken into account. The global convergence of temperature, excitation and ionisation structure requires solving Equations 5.2, 5.11, and 5.27 simultaneously, which represent a set of non-linear equations that are also solved by the Newton-Raphson method.
In this method, a starting guess for the global solution is made, where temperature and ionisation structure are arbitrarily chosen, and LTE excitation follows for that temperature, with the densities of the ejecta model at that epoch. The non-linear system of equations is linearised around the starting guess. This new linear system is then solved, and the procedure repeated around the new solution point. A requirement for convergence is that the linearised equations initially evolve in the same direction as the full non-linear ones, thus making good starting guesses essential.

Since KNe ejecta may have very different ejecta conditions than SNe, especially at later times due to increasing temperatures, the initial guess parameters typically used for the SN version of SUMO are not typically well suited for the later epochs of KNe. As such, initial guesses for the KN models are only used for the earliest epoch in the models (3 days in Paper II, 5 days in Papers I,III). These solutions are then used as the initial guess for the next time step, which is almost always a much better starting guess than the LTE solution for KNe past 5 days (as shown that LTE does not hold past 5 days in Paper II). I discuss the stability of the Newton-Raphson method, as well as step-size tests conducted during the thesis in Section 6.4 below.
6.2 Radiative Transfer

Radiative transfer represents a key part of this thesis, in particular for Paper III where KN spectral outputs of SUMO were presented for the first time. Although no modifications were made to the radiative transfer part of the code, it remains useful to describe the key functioning and processes modelled. The complete details are available in (Jerkstrand 2011; Jerkstrand et al. 2011, 2012). I described here in brief, more qualitative terms the key points of the RT modelling implemented in SUMO that differentiate it from other RT codes currently employed in the community, and discussed in Chapter 7.

From the solutions of the gas state calculated by solving the nebular phase equations described in Chapter 5, emissivities are calculated, for each radial zone and logarithmically spaced wavelength bin, with resolution $\Delta \lambda/\lambda = 0.001$ (see Section 6.4 below). From these emissivities, divisible photon packets are created, with energies sampled from a uniform distribution per zone and wavelength bin. The radius at which the MC packets are emitted are sampled from within the zone following a uniform distribution, and launched after sampling an isotropic distribution for the launch angle. These packets are then propagated through the medium, and continuously redshifted to account for the ejecta expansion, with matter interactions based on random draws.

Photon packets interact with matter broadly in two ways: by continuum processes, and by line interactions. Considering first continuum processes, these include photoionisation (PI), electron scattering (ES) and free-free interactions\(^1\). Each of these processes contributes an optical depth to the total optical depth of the region across which the photon packet travels. These processes are continuously applied across this travel distance, with the number of events for each process depending on its contribution to the total optical depth in that region. For instance, the number of photoionisations from level $l$ of atom $k$ in zone $i$ is given by:

$$\Delta N_{\text{PI}}^{i,k,l} = \frac{\tau_{\text{PI}}^{k,l}}{\tau_{\text{tot}}} \frac{E}{h \nu} (1 - e^{-\tau_{\text{tot}}})$$  \quad (6.1)$$

where $\tau_{\text{tot}}$ is the total accumulated optical depth in the region from all processes, including line interactions, $E$ is the packet energy, $\nu$ the packet frequency, and $\tau_{\text{PI}}^{k,l}$ is the optical depth for the ionisation of that specific level. The photoelectric heating rates are also

\(^1\)Dust absorption is included in SUMO but not for KN modelling

87
adjusted to take into account the contribution from these processes for the next calculation of temperature. The above number of absorbed photon packets is then used to find the PI rate for a given species using:

\[ \Gamma_{PI}^{i,k} = \frac{\sum_l N_{PI}^{i,k,l}}{V_i \sum_l n_{*,kl}^i} \]  

(6.2)

where \( V_i \) is the zone volume, and \( n_{*,kl}^i \) is the previous level population solution of level \( l \) from which the PI is occurring. We currently limit PI to the first 50 levels of the atom for memory and runtime purposes. Since these are typically the most populated levels, they are expected to largely dominate the PI flow, despite needing more energetic photons to be ionised. The above PI rate is then used when solving iteratively for the ionisation structure of the ejecta. The energy of the photon packet is attenuated by \( E_{new} = E_{e} e^{-\tau_{PI}} \), and is further propagated through the medium. Through this methodology, PI rates are therefore directly estimated from the MC packet interactions.

Line interactions occur when the photon packet reaches a line of sufficient optical depth to be included in the transfer calculations (currently \( \tau_{thresh} = 10^{-5} \), see Section 6.4 for discussion). If this occurs, then the total optical depth \( \tau_{tot} \) is increased by \( \tau_{line} \). SUMO considers two different treatments of line absorption depending on whether the upper level \( u \) is included in the fully coupled NLTE levels, or in a higher lying state above \( u_{max} \). Considering first the latter case, the photon packet will be completely absorbed if the added line optical depth makes \( \tau_{tot} > \tau_{abs} \), given by:

\[ \tau_{abs} = -\ln(1 - z) \]  

(6.3)

where \( z \) is a random number uniformly drawn between 0 and 1.

In this case, the deexcitation cascade will be calculated on-the-fly. All downward radiative and collisional rates are calculated following the descriptions in Chapter 5, and the radiative rates treated with the Sobolev escape probability formalism (Sobolev 1960; Jerkstrand 2011; Jerkstrand et al. 2011). The rates are normalised to obtain relative transition probabilities across all possible channels, though collisional excitation is omitted. If a radiative deexcitation is chosen, a new packet is emitted with an adjusted energy, and the process is repeated recursively until it and any potential offspring packets are absorbed or escape. Collisional deexcitations contribute to the volumetric heating rate. This on-the-fly treatment is an approximate treatment for the energy distribution of flu-
orescence when absorption occurs from an excited state, as the atom is forced to deexcite back to the ground state.

If the upper level is within those for which the full NLTE solutions are calculated i.e. \( u \leq u_{\text{max}} \), then the deexcitation cascade is not immediately calculated. The various radiative transition rates are adjusted based on estimators from the MC routine. The number of absorbed packets is increased for a transition from lower level \( l \) to upper level \( u \):

\[
\Delta N_{PE}^{i,k,l,u} = \frac{E}{h\nu} (1 - e^{-\tau_{l,u}^S})
\]  

(6.4)

The PE rate for a given transition between levels \( l \) and \( u \) is then estimated from the number of absorbed packets using:

\[
R_{i,k,l,u} = \frac{N_{i,k,l,u}}{n_{i,k,l}} \left( 1 - \frac{g_{k,u} n_{i,k,u}}{g_{k,l} n_{i,k,l}} \right)^{-1}
\]

(6.5)

The packet then continues in the same direction with an adjusted energy \( E_{\text{new}} = E e^{-\tau_{l,u}^S} \). Stimulated emission may occur for lines with \( \tau_{l,u}^S < 0 \), in which case the same formalism is used as for positive optical depths, and the deexcitation rate is found using:

\[
R_{i,k,u,l} = \left| \frac{N_{i,k,l,u}}{n_{i,k,u}} \right| \left( 1 - \frac{g_{k,u} n_{i,k,u}}{g_{k,l} n_{i,k,l}} \right)^{-1}
\]

(6.6)

The excitation rate is then found by using the relation \( R_{u,l} = R_{l,u} g_l / g_u \).

By using PE rate estimators from the number of absorbed MC packets, the emissivities calculated in the next iteration of the NLTE calculations take these radiative processes into account. The deexcitation channels available to photoexcited electrons are the same as in the on-the-fly approach, the difference now being that it is not immediately possible to disentangle whether excitation for a given electron was produced by collisional or radiative excitation. However, this approach is physically more accurate, and no assumption is made that an excited electron must deexcite back to the ground state. In this method, fluorescence and resonance scattering are therefore fully modelled with physical consistency, and are found to play key roles in KN spectral formation (Shingles et al. 2023, and Paper III). For r-process species, \( u_{\text{max}} = 1000 \), which is set to be identical to the maximum number of allowed levels in the atomic data, limited to this value due to memory limitations for computations. As such, every level considered by SUMO for r-process species
is treated in the second way. A discussion on the limitations of levels included can be found below in Section 6.4.

6.3 Atomic Data for Kilonova Modelling

The machinery that been described above for solving the ejecta gas state and generating the emergent spectrum must all naturally be applied to some sort of ejecta model, with a given elemental composition. While the composition may be user defined, and based on solar r-process abundances like in Papers I, II, or on the outputs of nuclear networks like in Paper III, the actual atomic data for these elements is still required. In concrete terms, this means having access to a set of levels and lines (transitions with wavelengths and probabilities) for the elements and their ionisation stages included in the model composition. The ionisation stages used thus far have been neutral to triply ionised, as motivated by various results for the ionisation structure of KNe in the NLTE regime (Hotokezaka et al. 2021, and Paper I). Such a feat is non-trivial however, as much of this data is experimentally, or even theoretically, unavailable for the heavy, r-process elements.

Due to such sparse availability, the RT community has worked hand in hand with the atomic physics community in order to calculate theoretical atomic data, using diverse atomic physics codes e.g. HULLAC (Bar-Shalom et al. 2001), GRASP2K (Jönsson et al. 2013), AUTOSTRUCTURE (Kasen et al. 2013), pseudo-relativistic Hartree-Fock method (Cowan 1981; Carvajal Gallego et al. 2022, 2023). The atomic data used in the context of this thesis was generated by collaborator Dr. Jon Grumer, using the Flexible Atomic Code (FAC (Gu 2008))\(^2\). This code uses a Dirack-Fock-Slater scheme for orbital optimization, and solves a Dirac-Coulomb Hamiltonian for structure calculations. Several levels of detail can be included with respect to electron interactions. In the most recent data set employed for Paper III, Breit interactions in the low-frequency limit were included, and leading order quantum electrodynamical effects, such as vacuum polarisation and electron self-energy were also considered. The FAC methodology takes a set of configuration states as input, and works on the average configuration to optimise a common central potential. This type of model is referred to as a spectroscopic configuration-interaction model, which includes only the most fundamental many-electron effects on the energy levels and transitions.

\(^2\)FAC is available at source available at https://github.com/flexible-atomic-code/fac.
For Papers I and II, a set of atomic levels and lines were generated for the neutral to triply ionised ions of Te, Ce, Pt and Th. For Paper III, a complete set for the same ionisation stages, for all elements from Cu to U were generated, representing one of the most complete atomic data sets to date in the field. Notably, the transitions included forbidden electric quadrupole (E2) and magnetic dipole (M1) transitions typically omitted by other data sets, but found to play significant roles in determining the ejecta gas state (see e.g. Papers I and III).

When generating r-process atomic data for spectral modelling, current computational limitations typically force one to decide whether an accurate, or complete data set is desired. In the context of this PhD thesis, where the broader ejecta conditions and emergent spectral shapes were considered, a complete data set was preferable. Other similar data sets were used to calculated expansion opacities broadly employed by the KN RT community (Tanaka et al. 2020), and it was found that opacities from our data set were consistent with literature values (see Paper II). This provides support in the completeness and consistency of the atomic data employed in this thesis, with respect to well established and accepted data sets in the community.

Naturally, such a complete data set cannot be fully accurate with respect to the experimental atomic data that does exist. Notably, the energies of the atomic levels and corresponding transition wavelengths are often shifted relative to the ‘established’ energies found in databases such as the NIST Atomic Spectra Database (Kramida et al. 2020), Vienna Atomic Line Database (VALD3, Ryabchikova et al. 2015), and DREAM (Biémont et al. 1999), which are typically compiled from a mixture of experimental measurements and best theoretical calculations. The transition strengths, or A-values, are hard to gauge in terms of accuracy, due to the lack of data for anything but the strongest allowed transitions. Such databases cannot be used directly in numerical simulations like those presented in this PhD thesis, as they are not complete enough in data, especially for transition strengths and generally for IR transitions. While data may be abundantly available for certain elements like the first r-process peak elements observed in stellar spectra, or fusion related elements like highly ionised tungsten, data is largely incomplete for elements like the lanthanides, which are found to play key roles in KNe (e.g. as shown in Paper III).

Shifted energy levels and wavelengths may be particularly important for species with strong allowed transitions which have significant impacts on the spectral formation, such
as the Sr II 10000Å triplet. In such cases, it is useful to shift the theoretical levels to match those in experimental databases, as was done for Sr II in Paper III. However, this is not often trivial, as the nature of the level labelling depends on the electron coupling formalism used, which often differs between theoretical codes and experimental databases. This makes identifying which theoretical levels correspond to which experimentally measured levels difficult, and near impossible for all but the simplest atomic structures, e.g. single valence electron species like Sr II.

As such, the spectral predictions of Paper III were presented in the context of broad SED shape and evolution, as opposed to the detailed analysis of individual features. The future of atomic data likely now lies in calibrating or improving the accuracy of these large, complete data sets, firstly for key species like those identified in Paper III, and then for less significant species down the line, for completeness. For Papers I and II, the statistical completeness of the levels and transitions was more important, in order to determine the overall gas state of the ejecta, and so no rescaling of atomic data was necessary.

6.4 Numerical Tests and Code Limitations

As with any numerical simulations including complex physics, treatment of various processes are often simplified using diverse assumptions. The numerical techniques themselves also have limitations in their application, which may affect the accuracy or stability of the solutions. It is therefore good practice to verify the impact of such limitations, as well as those arising from choices made by the user. In this section, I discuss some of the testing that was conducted throughout this thesis in order to assess the reliability of the KN simulations conducted with SUMO.

6.4.1 NLTE Solutions: The Newton-Raphson Method

As mentioned previously in Section 6.1, the NLTE rate equations are solved using an iterative Newton-Raphson scheme. The stability of the Newton-Raphson method depends on the step size used between iterations, where large steps decrease stability. This means that it is often preferable to use a damped step size as opposed to a full one, in order to reduce the risk of overshooting the solution to a non-convergent regime. As such, we currently employ a damped step size of $h = 0.8$ for temperature, $h = 0.9$ for excitation,
and \( h = 0.5 \) for ionisation, where \( h = 1 \) is a normal step size. We also limit how much the temperature solution can change in a single global iteration, likewise to reduce the risk of overshooting into a non-convergent regime, such that the new temperature solution must be within 20\% of the old one: \( 0.8 T_{\text{old}} \leq T_{\text{new}} \leq 1.2 T_{\text{old}} \). While this may increase the runtime of simulations, especially when the starting guess is far from the true solution, it reduces the risk of converging to the incorrect solution by overshooting.

Determining convergence can be tricky, and various checks are put in place to ensure that a properly converged solution is found. For temperature, we require \( T_{\text{new}} / T_{\text{old}} < 10^{-3} \) between consecutive Newton-Raphson steps, and the difference between overall cooling and heating rates to be \( \leq 1\% \) of the heating rate for that solution. Should these conditions not be reached within 15 iterations (e.g. due to the maximum temperature change being reached for a single global \textsc{SUMO} iteration), then the solution from the final Newton-Raphson iteration is used. This typically only occurs during the first few global iterations from an initial guess. For excitation structure, we require \( n_{\text{new}} / n_{\text{old}} \leq 0.01 \), while ionisation structure requires \( x_{\text{new}} / x_{\text{old}} \leq 0.005 \). Ionisation convergence in steady-state also requires that the sum of net rates (ionisation - recombination) squared is smaller than \( 10^{-8} \).

The global convergence initially relied on the overall energy balance (EB) in the simulation. This is the ratio of energy going into the simulation from radioactive energy deposition, to all processes emitting radiation. This convergence criteria is well adapted to the nebular phase, where the emergent luminosity is expected to track the instantaneous energy deposition. For Papers I and II, global convergence in steady-state mode required that the energy balance be around 1.0, with a tolerance of 0.05, such that values \( 0.95 \leq EB \leq 1.05 \) were allowed. A minimum of 30 global iterations were required before the convergence was accepted, with checks in place if the energy balance was going into non-convergent regimes (e.g. if \( EB > 3.0 \) for more than 20 iterations consecutively).

In time-dependent mode, processes like adiabatic cooling may lead to energy balance dropping below 1.0, and the emergent luminosity no longer tracking the instantaneous deposition (see Figure 6 in Paper I). In Paper I, convergence in time-dependent mode was taken by forcing the model at a given epoch to go through all 50 maximum global iterations, even though most of the models in this mode still had energy balances within the steady-state convergence tolerance. For the simple 4-element compositions of Paper I, it was found that the models tended to converge within 50 iterations when taking the
Figure 6.2: Convergence of bolometric luminosity and energy balance (top panel), temperature solutions (middle panel) and ionisation solutions/electron fractions $x_e$ (bottom panel) for the $Y_e \sim 0.35$ model of Paper III at 5 days. The initial guess was the default used in SUMO, which is quite far from the true KN solutions. Own figure.
default initial guess that was typically far from the real solution, so 50 was chosen as an upper limit to avoid needlessly using too much runtime. Furthermore, the more complex, 30-element models were also found to have the gas state practically converged within 50 iterations, as shown in Figure 6.2. Given that the Paper I models converged much earlier and did not move away from the converged solutions with further iterations, 50 maximum iterations was assessed to be a safe upper limit.

In Paper III, the energy balance convergence was abandoned, as it was found that the adiabatic degradation of the radiation field due to many scattering interactions at early times was reducing the final energy balance by $\sim 5\% - 20\%$ (see Paper III), and thus beyond the tolerance of 5% energy balance used in steady-state mode. As such, convergence criteria were implemented based on maximum allowed changes in temperature, ionisation fraction and bolometric luminosity between consecutive global iterations. For temperature and ionisation fraction, a maximum change of $x_{\text{new}}/x_{\text{old}} \leq 0.1\%$ in any zone was set, while a $L_{\text{new}}/L_{\text{old}} \leq 1\%$ tolerance was set for the bolometric luminosity. The higher tolerance for luminosity variations was adopted after finding that this value tended to bounce around the solution with changes on the order of $\sim 0.5\%$ between consecutive iterations, notably in the heavier composition model of Paper III. The evolution of the thermodynamic properties from initial guess to converged solutions are illustrated for the $Y_e \sim 0.35$ model from Paper III at 5 days in Figure 6.2.

### 6.4.2 Atomic Data Limitations

Another key aspect to consider in the convergence of the solutions to the true values is how complete the atomic data (levels and lines) must be in order to find the real solution. Due to the enormous complexity of certain r-process species, notably the lanthanides, not all the levels available in the provided theoretical atomic data could be included. Indeed, certain lanthanide species, such as neutral neodymium (Nd I) have a theoretical level count on the order of $\gtrsim 3400$ levels, which cannot all feasibly be included in the NLTE calculations. Even reading in so many levels and having the majority of them not fully coupled to the NLTE solutions, such that the on-the-fly cascade method would be used above in the radiative transfer, is not feasible in terms of memory usage. This is without even considering the corresponding millions of associated lines. As such, for r-process species, a maximum atomic data level count of $n_{\text{max}} = 1000$ was implemented. This was
also the value taken for the maximum upper level to be fully coupled to the radiative transfer, \( u_{\text{max}} \), as described above in Section 6.2, such that all 1000 levels were treated fully coupled in the NLTE solutions.

The value of \( n_{\text{max}} = 1000 \) was chosen by checking how many levels could be included in the heaviest composition model of Paper III with \( Y_e \sim 0.15 \), without exceeding the available cluster memory. This heavy model included almost every lanthanide species in the available atomic data, as well as the actinides thorium (Th) and uranium (U), all of which are incredibly level and line rich. There are several caveats to limiting the level count in such a way. The first is that the entire range of excitation energies for a given species is not fully sampled, as the lowest 1000 lying states are taken. The extent to which this matters will depend on how much of the energy range is sampled in the first 1000 levels. In the case of Nd I, which is the largest species of our sample with 3405 levels, the lowest lying 1000 states reach an energy value of \( \sim 30000 \text{ cm}^{-1} \), with the ionisation limit being \( \sim 44500 \text{ cm}^{-1} \) (Martin 1978). As such, approximately 33\% of the excitation energies available are not considered using this cut-off value. However, the highest-lying states of any species are expected to be significantly depopulated compared to lower lying states, especially past the first few hundred levels, as was shown in Paper II. Therefore, it is likely that these states do not play a large role in the gas state solutions or emergent spectra.

Another limitation that may be implemented is on the number of transitions included in the atomic data. The atomic data that is provided includes allowed E1 dipole transitions, as well as forbidden electric quadrupole E2, and magnetic dipole M1 transitions. In cases where there are many thousands of levels for a species, the number of included transitions is far greater. In the example of Nd I above, a total of over 2.9 million transitions were included in the data, which is not a feasible number to treat in a line-by-line method. In many physical contexts however, it is the stronger lines that dominate the spectral formation, and play the largest roles in the thermodynamic solutions. As such, one may consider cutting out weaker transitions by using a branching ratio cut. We test here whether a branching ratio cut can be safely implemented without changing the emergent solutions. Ratio cuts of \( 10^{-6} \) and \( 10^{-3} \) were tested on the \( Y_e \sim 0.25 \) model from Paper III at 5 days, the results are shown in Figure 6.3. While the effect is not major for this model and epoch, excessive branching ratio cuts at later times, when forbidden transitions begin to emit significantly, may have larger impacts.
Figure 6.3: Tests of different branching ratio cuts on the $Y_e \sim 0.25$ model of Paper III at 5 days. We see that while a cut of $10^{-6}$ reproduces approximately the same spectrum, a cut of $10^{-3}$ changes the emergent spectrum, though individual features from strong transitions remain the same. Own figure.
Another way to affect the lines treated by SUMO is to change the value of $\tau_{\text{thresh}}$, which sets the minimal optical depth for a line to be treated in the radiative transfer. This is different from cutting out transitions from the raw data as above, since the optically thin transitions still exist and provide (de)excitation channels for the electrons. So while this does not reduce the time needed to read in and store numbers from the raw atomic data, increasing $\tau_{\text{thresh}}$ may speed up the radiative transfer portion of the simulation. However, a too high value will lead to issues with convergence, as the emergent bolometric luminosity will be mismatched with the incoming energy deposition. Since the ejecta conditions and emergent spectra are intrinsically and complexly linked in NLTE, omission of too many lines may also change the thermodynamic solutions.

The default value used for the SUMO KN simulations is $\tau_{\text{thresh}} = 10^{-5}$. For the $Y_e \sim 0.15$ model of paper III, this value was raised to $5 \times 10^{-5}$ at 7 days, and $10^{-4}$ at 5 days. The higher density at these early times lead the number of lines per zone in this lanthanide-rich model to exceed the maximum limit of 35000, which was chosen as the maximum value that did not exceed available memory. The effect of doing so at these early times is likely not too severe, as strong, optically thick allowed lines are expected to have the greater effect on the radiation field. I show here a test of $\tau_{\text{thresh}} = 10^{-6}, 10^{-5}, 10^{-3}$ on the solutions of the lighter $Y_e \sim 0.35$ model at 10 days, in Figure 6.4, to illustrate the effect in a context where memory usage is not a limitation. In the case of $\tau_{\text{thresh}} = 10^{-3}$, convergence was not properly achieved, and the energy balance was stuck around $\sim 72\%$, compared to $\sim 86\%$ for the lower values, resulting in a lower bolometric luminosity. Qualitatively, the spectral shape remains the same however, since it is mostly allowed transitions shaping this spectrum.

6.4.3 Resolutions and Ranges

A key aspect of any numerical simulation is the range over which quantities are considered, and the resolution with which that range is resolved. In the case of SUMO, the spatial, wavelength, and time dimensions are the main aspects to investigate. Let us consider first the spatial dimension, which here is a matter of radial zones. In Papers I and II, simple single-zone toy models were used. However, in Paper III, 5 zone models were employed, ranging over a velocity range of $0.05 - 0.3c$, increasing linearly in steps of $0.05c$. The total ejecta mass in all papers has been $0.05M_\odot$, and a $\rho \propto v^{-4}$ density profile was used for the
Figure 6.4: The effect of changing $\tau_{\text{thresh}}$ on the emergent spectra of the $Y_e \sim 0.35$ model of Paper III at 10 days. Note that the $\tau_{\text{thresh}}$ model struggles to converge and has an abnormally low EB and bolometric luminosity. Own figure.
models in Paper III, as motivated by relativistic hydrodynamical models of BNS mergers (e.g. Kawaguchi et al. 2021).

Before deciding on the number of zones, or alternatively, spatial resolution of the models, sensitivity tests were made. I show an example of a spectral output test for a 30 element, solar composition model at 20 days after merger in Figure 6.5. It was found that 5 zones corresponded to a good balance of solution convergence and computational power, since having a larger zone count requires a larger number of cores to run on, as well as longer run-times. Though there is some slight variance in individual features, the intrinsic inaccuracy of the atomic data itself, as well as approximate treatment of many processes, far exceeds the possible inaccuracies introduced from using 5 zones as opposed to a more resolved 10 zone model.

Another aspect to consider is the wavelength range and spectral resolution to use for the simulations. For the KN models, the spectral range of wavelengths was set from
Figure 6.6: Testing the effect of spectral resolution on emergent spectra. These are identical one-zone models with solar abundance compositions of Te, Ce, Pt and Th, as used in Paper II. The spectra have been smoothed by 2000 km s\(^{-1}\) in order to reduce MC noise. Based on these results, we choose $\Delta \lambda_{1000} = 1$. Own figure.

500Å to 50000Å. The lower limit corresponds was originally defined in the context of SNe, by considering a balance between reduced UV emissivities, and proper inclusion of photoionising photons for the temperature solutions relevant to nebular phase SNe ($T \lesssim 10000$ K) and the ionisation potentials of light elements. In KNe, temperatures may be much hotter than in SNe, but the typical ionisation potential of heavy elements is lower. These will lead to contrasting effects, which were not investigated in this thesis, and so the original lower wavelength limit was kept. The upper limit corresponds to the NIR at a wavelength slightly longer than the last AT2017gfo \textit{Spitzer} detection in the 3.4 $\mu$m band and non-detection in the 4.5 $\mu$m band (Kasliwal et al. 2017). The NIR wavelengths are particularly important, as many lines from the lanthanides and actinides are present there.

Tests on spectral resolution were conducted in the context of Paper II, to check how variable the converged solutions were between iterations. The spectral resolution is defined
Figure 6.7: Comparison of thin and full spectra for two models used in Paper III. In this case, the full radiative transfer was extended to 30 micron, to gauge whether transfer effects were ongoing in the MIR. The $Y_e \sim 0.15$ model could not be run earlier with such a spectral range, as this caused the number of optically thick lines to exceed the maximum limit. Own figure.
on a logarithmically spaced wavelength array with $\Delta \lambda / \lambda = \text{constant}$. We tested resolution values at $\lambda = 1000$ Å, given by $\Delta \lambda_{1000} = 0.5, 1.0, 10.0$, for a fixed total photon packet count checking the variation in spectral output between two consequent, converged iterations (for a single resolution), as shown in Figure 6.6. We found that a resolution of $\Delta \lambda / \lambda = 0.001$ appeared to be the best suited for our purposes, providing a balance between stability, accuracy, and runtime.

The cutoff for radiative transfer being set to 50000 Å was also motivated by the fact that reduced transfer effects are expected in the MIR, due to greatly reduced optical depths. This can be seen by comparing the initial ‘thin’ spectrum, e.g. the emissivities before the radiative transfer, to the ‘full’ spectrum obtained after the radiative transfer step. As shown in Figure 6.7, the amount of transfer occurring in the MIR can vary significantly depending on epoch and model. Notably, ejecta with significant abundances of IR emitting r-process elements like the lanthanides will likely still have a significant amount of transfer in the MIR at early times. Since the focus of Paper III was the UV/optical/NIR regimes, the cut-off of 50000 Å is justified, but this limit would likely have to be extended to 30 micron in the context of a JWST study. Depending on epoch and model composition, this may force lines to be removed from the transfer, e.g. by raising the $\tau_{\text{thresh}}$ value, in order to not exceed the maximum line limit per zone. The runtime is not expected to be significantly affected however, as the MIR regime should still be more optically thin than the UV/opt/NIR regardless.

With respect to time-step resolution, this only plays a role if the simulations are run in time-dependent mode. There, the previous solution directly affects the solution at the current epoch. As such, the time-step between epochs should not be too large to prevent entering a non-convergent regime. We take a time-step of 10% between epochs, having also tested a 20% time-step and finding that the temperature and ionisation structure solutions are practically identical and well converged in both cases, thus implying the 10% time-step to be acceptable (see Appendix D in Paper I).
7 Kilonova Modelling Techniques

In this Chapter, I cover how diverse radiative transfer (RT) codes simulate KNe, and compare these to SUMO. These codes vary in their approaches and numerical techniques, but all commonly employ a Monte Carlo (MC) radiative transfer method. The treatments of the ejecta thermodynamics, as well as details pertaining to the RT methodology itself are more variable. The following codes will be considered: SEDONA (Kasen et al. 2006, 2013, 2015), ARTIS (Kromer & Sim 2009) and the simplified version TARDIS (Kerzendorf & Sim 2014), and a code initially developed by Tanaka & Hotokezaka (2013), henceforth referred to as TH13. I also consider the methodology employed in (Hotokezaka et al. 2021), which remains the closest study to full NLTE modelling like SUMO so far. Throughout this section, these codes will be compared to each other, but also to the functioning of SUMO, which is the code used for this PhD thesis.

Aside from the codes mentioned above, that are covered in more detail, it is worth mentioning two more codes that have been used by the community: SuperNu (Wollaeger et al. 2013; Wollaeger & van Rossum 2014), and POSSIS (Bulla 2019). SuperNu is a 3D, LTE code employing expansion opacities that is particular in the way it treats the MC approach, by using a hybrid method making use of both implicit and discrete diffusion MC. Implicit MC allows for solving time-dependent RT coupled non-linearly to matter in a stochastic manner, but suffers from long run-times when modelling optically thick regimes with many scattering interactions. Discrete diffusion MC treats this diffusion regime by approaching the diffusion equation again in a stochastic manner. Hybridisation of the two methods therefore allows for efficient treatment of the MC method even in optically thick regions. SuperNu has been used in several studies to calculated synthetic LCs and spectra for various ejecta morphologies (Wollaeger et al. 2018, 2019, 2021).

POSSIS is a 3D, time-dependent MC code which stands out by also predicting polarisation of the emergent light. As for many codes, it operates using LTE physics and
expansion opacities to conduct the radiative transfer, where opacities are either treated
under the Sobolev approximation, or with a polynomial fit to the line opacities which are
used to predict a featureless 'pseudo-continuum' spectrum (Inserra et al. 2016). Polari-
sation arises from electron scattering events, while other matter interactions are treated
using the ‘Two Level Atom’ (TLA) approach described below in Section 7.3. The emerg-
ent spectra are either found by collecting escaping photons, as in most MC codes, or by
collecting virtual photon packets generated after MC packet matter interactions, in what
is referred to as an event-based technique. These virtual packets are weighted according
to their probability of reaching the observer after their generation, and allows synthetic
observables to be created with reduced MC noise (Bulla et al. 2015). POSSIS has been
applied to various KN studies (e.g. Bulla 2019; Bulla et al. 2021), with further applica-
tions of using KNe to measure the Hubble constant (Coughlin et al. 2020; Bulla et al. 2022), as
well as investigations into the impacts of jets on KN emission (Nativi et al. 2021; Shrestha
et al. 2023).

7.1 TARDIS

TARDIS is a 1D, LTE code, which assumes a spherically symmetric explosion with ho-
mologously expanding ejecta, the density profile and composition of which is set by the
user. TARDIS has also recently been upgraded to fully treat special relativistic effect (Vogl
et al. 2019). An inner boundary with properties (e.g. velocity, temperature) defined by
user input is used as a photosphere, with material assumed to be optically thick below
this boundary. This photosphere emits as a perfect blackbody (BB) into line-forming
regions placed above it (see Figure 7.1). As such, TARDIS is well suited for modelling for
the photospheric phase of the KN. It is assumed that the material in the computational
domain is in radiative equilibrium, and that the only energy source is radiative (e.g. no
radioactivity). TARDIS may solve for temperature, ionisation, excitation and conduct the
radiative transfer component in different modes of varying complexity and accuracy, origi-
nally defined in ARTIS. Kilonova studies with TARDIS have thus far limited themselves in
which options are used, so I only describe those relevant for current existing studies. An
example of a TARDIS model spectrum fit to AT2017gfo can be seen in Figure 7.2.

Indivisible photon packets are generated by sampling the BB continuum, and propa-
gated though the line-forming regions. The initial conditions of these outer regions, such
as density, composition and temperature are determined by user inputs. Photon packets may interact with matter by free electron scattering, bound-bound, and bound-free/free-bound processes. Gas temperature in TARDIS is typically found by enforcing radiative equilibrium, which amounts to solving for radiation temperature in each grid cell, such that the following equation is satisfied:

\[ J_\nu = W B_\nu(T_R) \]  

(7.1)

where \( J_\nu \) is the frequency dependent mean intensity, \( W \) the dilution factor, \( B_\nu \) the Planck function, and \( T_R \) the radiation temperature. The electron temperature is taken to be \( T_e = 0.9T_R \) (Kerzendorf & Sim 2014). Many studies have however opted to employ user defined fixed ejecta temperatures instead (Smartt et al. 2017; Watson et al. 2019; Gillanders et al. 2022), motivated by poorly known opacity distributions and extreme UV line-blanketing, which may make the usage of a local blackbody radiation field unreliable due to flux redistribution (see Gillanders et al. 2022, and as found in Paper III).

The ionisation structure has thus far been calculated with the LTE mode, using the Saha equation, with the exception of Perego et al. (2022) that used the ‘nebular’ approx-
Figure 7.2: Best fitting TARDIS spectral models to the 1.4 and 3.4 day epochs of AT2017gfo, varied by electron fraction $Y_e$. Models 'a' and 'b' with the same $Y_e$ have similar compositions, with a few differences arising from differences in the nuclear networks used to generate the composition. From this kind of analysis, the lanthanide fraction of AT2017gfo was constrained to be $X_{lanth} \sim 5 \times 10^{-2}$, and two ejecta components were inferred, supporting the two-component ejecta models. Taken from Gillanders et al. (2022).
imation equation (see Equation 3 of Kerzendorf & Sim 2014). The excitation structure may be calculated in LTE from the Boltzmann equation using $T_R$ (used by Watson et al. 2019), 'dilute-LTE' where the BB dilution factor $W$ is taken into account, calculated iteratively from MC packet estimators (used by Gillanders et al. 2021, 2022), and a 'NLTE' option which calculates the level populations from rate equations and thus far used in a KN context by Perego et al. (2022).

Radiative rates are used when the 'nebular' and 'NLTE' modes are activated, and then either estimated using the 'dilute-blackbody' method using Equation 7.1 for a given transition, with values coming from MC estimators, or by a 'detailed' treatment that uses MC estimators directly without enforcing Equation 7.1. An additional estimator for the intensity in the blue wing of a given line is also calculated in this treatment. Line interactions may also be treated in various ways, the simplest being by assuming pure scattering. The most complex that has been used in TARDIS is the full 'macro-atom' treatment (Gillanders et al. 2021), which was originally developed for ARTIS and is described below. The key advantage of this method is that fluorescence, which was found to play a key role in KN spectral formation (Shingles et al. 2023, and Paper III), is fully modelled.

Due to its flexibility of operating modes, TARDIS has provided some successful interpretations of the AT2017gfo data, with varying levels of detailed being considered, some of which, such as the macro-atom, approach SUMO in terms of level of detail. The combination of relatively simple ejecta and photospheric set-up that leads to fast run-times, combined with the possibility for detailed radiative transfer by usage of the macro-atom method, maintains TARDIS as a leading tool for rapid KN analyses to this day.

7.2 ARTIS

ARTIS works in a relatively unique fashion, in the sense that it was originally developed with a particular methodology called the macro-atom (Lucy 2002, 2003, 2005). The macro-atom approach makes use of various different indivisible 'packets': kinetic thermal energy 'k-packets', radiation field 'r-packets', and internal energy 'i-packets'. This method is based on a macroscopic interpretation of radiative and statistical equilibrium considering energy flow rates between the thermal kinetic pool, the radiation field, and the levels of the species defined as macro-atom states. Of the defined packets, only the r-packets are
propagated and interact with matter using a MC method.

The macro-atom method considers all possible excited levels of a species as energy pools, in and out of which flows energy by available transitions. Statistical equilibrium is satisfied by imposing a set of flow rules between states, which drive internal (non-radiative/collisionless) transitions between levels of the macro-atom. Energy flows from kinetic or radiative processes are implemented in a MC context by corresponding to the generation or destruction of k-packets or r-packets. For example, when PE occurs, an r-packet is absorbed, with the relevant absorption process defining which macro-atom state is ‘activated’ and an i-packet of equal energy is generated. For collisional excitation, a k-packet is destroyed, and the activated state is determined by sampling cooling rates for collisional excitation, with an equal energy i-packet again being generated.

When a macro-atom state is selected, or activated, all rates for transitions connecting that state to other states are calculated, as well as rates converting the i-packet back to thermal kinetic or radiation energy, which is considered to ‘deactivate’ the macro-atom state. A transition is randomly selected from all the possibilities by probability sampling, with the process being repeated until a deactivating process is selected. When this occurs, the co-moving frame energy of the packet is conserved, which assures radiative equilibrium. Collisional deexcitations convert the i-packet into an equal energy k-packet, while radiative (de)excitations convert to an isotropically emitted r-packet. The associated frequency of the r-packet is either the transition frequency for bound-bound transitions, or sampled from the emissivity for free-bound transitions. This method is conceptually similar to the ‘on-the-fly’ cascade approach used by SUMO for partially coupled NLTE energy levels, and as such allows for a fully consistent treatment of radiation matter interactions, including scattering and fluorescence. In this sense, ARTIS is the only code used for KN modelling that may, in theory, be used to model KNe with a NLTE line-by-line RT approach, though it has so far only been used in the LTE framework (Collins et al. 2023; Shingles et al. 2023).

In the original ARTIS code described in (Kromer & Sim 2009), the energy deposition originated purely from $\gamma$-ray decay, with decay energy being placed in indivisible $\gamma$-packets and propagated through the ejecta using wavelength-dependent MC transfer. These then interact with matter by Compton scattering, pair production and photo-absorption, the latter of which transforms them into thermal kinetic energy k-packets. The k-packets are transformed by various processes into UVOIR radiation r-packets, and internal energy i-
packets following the macro-atom method, with both of these packets also being converted into the other types respectively depending on the process sampled. Propagation of r-packets follows the MC scheme in a similar fashion to SUMO, by considering the distance to the nearest critical point, defined as the edge of the grid, the end of the time-step, or matter interaction, where the latter is determined when the accumulated optical depth over the travel distance transcends a randomly drawn depth (Equation 6.3). The Sobolev approximation is used for line interactions, and continuum interactions are decided by sampling the absorption coefficients of the various processes.

ARTIS has currently been used in the KN context for only a couple studies (Collins et al. 2023; Shingles et al. 2023). These both used the same low-mass (0.005\(M_\odot\)) dynamical ejecta hydrodynamical model, with atomic data coming from various sources (e.g. Hillier \& Miller 1998; Kurucz 2018; Tanaka et al. 2020). In both cases, the total energy of the simulation and its distribution is defined at the start of the simulation, according to the total energy released by the smoothed particle hydrodynamics particle trajectories in the model grid cell. The energy deposition rate, corresponding to the time when an MC packet is placed in the cell, is also defined at the start. This is assumed to be identical
for all grid cells, and is taken as the average raw radioactive power $\dot{Q}$ of all unbound SPH particle trajectories (Collins et al. 2023). Packet energy loss from adiabatic degradation as time goes on is taken into account.

Wavelength-dependent $\gamma$-ray transport is fully modelled, and Shingles et al. (2023) added a detailed time-dependent thermalisation treatment for $\beta$-decay electrons and $\alpha$-particles, while Collins et al. (2023) used an averaged $\gamma$-ray emission spectrum and assumed instant thermalisation of $\beta$-decay electrons, omitting $\alpha$-decay products. The full energy deposition scheme calculates the integrated energy release in a given cell at a given time, and is used to seed the thermal energy k-packets. While Collins et al. (2023) used a grey opacity treatment (Sim 2007), Shingles et al. (2023) conduct a line-by-line transfer treatment within the Sobolev expansion formalism. This allowed them to determine that fluorescence and scattering play key roles in the early photospheric LTE phase of KNe, a result that is echoed in Paper III of this thesis. Their model broadly reproduced the spectral shape of AT2017gfo as early times, as shown in Figure 7.3.

7.3 SEDONA

SEDONA is a 3D, time-dependent MC transfer code originally designed in the context of SNe (Kasen et al. 2006). It is flexible in terms of allowed numerical grids, including 1D spherical, 2D cylindrical, and 3D Cartesian coordinate systems. Density, composition and any shock heating deposition are specified at the initial time of the simulation, with 3D wavelength-dependent $\gamma$-ray transport also being carried out to find the energy deposition in each cell. Adiabatic losses of the radiation field are modelled, but relativistic corrections on the order of $(v/c)^2$ are neglected. The ionisation and excitation structures are calculated using the LTE assumption and using the Saha-Boltzmann equations, and radiative equilibrium is enforced. Scattering and fluorescence terms are included in the line source function, where fluorescence is approximated by a single step de-excitation along a different channel than the original excitation channel, i.e. the full cascade is not followed.

SEDONA calculates temperature by converging temperature and wavelength dependent (from LTE excitation/ionisation structure) opacities and emissivities (Mihalas 1978; Kasen et al. 2006), with the temperature solution found from setting the rate of thermal electromagnetic emission equal to the rates of photon absorption plus radioactive energy deposition (e.g. Kasen & Barnes 2019, for energy deposition). The radiative transfer is
conducted using the Sobolev approximation, with electron scattering, free-free, bound-free processes modelled using standard opacity formulae, with the hydrogenic formulae being applied otherwise.

For line interactions, three processes are considered with respect to packet redirection and frequency redistribution: pure scattering, thermal absorption and re-emission from thermal sampling, and approximate fluorescence. A subset of the most important lines are selected to be treated with a direct individual (e.g. line-by-line) treatment and approximate fluorescence, while the rest of the lines are binned using the expansion opacity formalism (Karp et al. 1977; Eastman & Pinto 1993). These lines are treated using the ‘Two Level Atom’ (TLA) method, where the source function is defined as:

\[ S_\lambda = (1 - \epsilon)J_\lambda + \epsilon B_\lambda(T) \]  (7.2)

where \( J_\lambda \) is the radiation field mean intensity, \( B_\lambda \) is the Planck function, \( \epsilon \) is the probability of absorption, generally unique for each transition. A common value of \( \epsilon \) is typically chosen for every line in order to simplify the TLA treatment.

Propagation of the MC photon packets works in a similar fashion as to SUMO, with calculation of distances to grid cell edges and matter interactions. Escaping photon packets are collected and binned in order to form the emergent spectrum. Continuum interactions occur after the photon has moved a corresponding velocity distance of:

\[ v_C = -\frac{1}{\alpha_{\text{cont}} \text{exp} \log(z)} \]  (7.3)

where \( \alpha_{\text{cont}} \) is the total opacity (sum of free-free, bound-free, electron scattering and line expansion opacities), and \( z \) is a random number between 0 and 1. Continuum interactions may be absorption, electron scattering or expansion-opacity line scattering. Absorbed packets are assumed to immediately thermalise, and are instantaneously re-emitted after sampling from a thermal distribution. Interaction with a line treated directly occurs if the condition \( z < 1 - e^{-\tau_{\text{TLA}}} \) is met, conceptually similar to the condition for on-the-fly line interactions in SUMO. The packet may then be absorbed or radiatively de-excite, including by single-step fluorescence. Polarisation of the radiation field may also be calculated.

SEDONA has been used for several KN studies (Kasen et al. 2013; Barnes & Kasen 2013; Kasen et al. 2015, 2017), in which all lines were treated using the expansion opacity method. As such, line interactions were treated using the simplified TLA approach. In
Figure 7.4: SEDONA models of spectra and multi-band LCs for a 2-component KN model, used to explain the broad ejecta properties of AT2017gfo. The red, lanthanide-rich component grows is slower moving and more massive than the blue component, and so grows to dominate the later time emission while the blue component fades away. Taken from Kasen et al. (2017).

this method, fluorescence is not modelled, and photon packets undergoing line interactions may either scatter or be absorbed and thermally re-emitted depending on the redistribution probability parameter $\epsilon$ as defined in Equation 7.2. In this method, absorption by a line is immediately followed by re-emission in another line according to a thermal distribution, which 'mimics' fluorescence. The sensitivity of the TLA method in reproducing full fluorescence treatment SN spectra is explored in Kasen et al. (2006). In KN studies making use of SEDONA, the choice of redistribution parameter has thus far assumed purely absorptive lines, such that $\epsilon = 1$, with the epochs studies typically being limited to early time, photospheric phase LCs and spectra, as shown in Figure 7.4.

### 7.4 TH13

The TH13 code is likewise a 3D, time-dependent MC code, that was developed with both SNe and KNe in mind (Tanaka & Hotokezaka 2013), and has been readily applied to KNe in the past few years (e.g. Tanaka et al. 2017, 2018, 2020; Kawaguchi et al. 2021; Domoto 2020).
et al. 2021, 2022; Kawaguchi et al. 2023). Similarly to methods described above, an ejecta morphology with composition and density is taken as input. γ-ray packets are generated for SN cases and propagated using a grey approximation (wavelength-independent) transfer. Absorbed γ-ray packets are converted into indivisible radiation packets. For KNe, the packets have an activation lifetime such that the energy generation rate follows the well established $\dot{E}_{\text{decay}} \sim t^{-1.2}$ power law for ensemble β-decay of r-process nuclei. In the original code, a simplified thermalisation form of $\dot{E}_{\text{rad}} = (0.1 + 0.9\epsilon_{\text{therm}})\dot{E}_{\text{decay}}$, where $\epsilon_{\text{therm}} = 0.5$ was chosen, such that the energy $\dot{E}_{\text{rad}}$ was immediately deposited without transportation (Tanaka & Hotokezaka 2013). Later works either used analytical formulae from Barnes et al. (2016); Kasen & Barnes (2019), or derived their own thermalisation efficiencies (Hotokezaka & Nakar 2020). At the activation time, UVOIR packets are directly created (instead of γ-packets), with the initial co-moving wavelength assigned by sampling the emissivity $j_v$.

The UVOIR packets are propagated using the usual MC techniques including opacity from electron scattering, free-free, bound-free and bound-bound interactions. A critical distance is calculated to the next scattering/absorption event, edge of the cell, and end of time-step, with the first option being based on a threshold optical depth $\tau = -\ln(z)$. Line interactions are treated using the Sobolev approximation and expansion opacity formalism, with lines being purely absorptive, corresponding to the $\epsilon = 1$ treatment as described above for SEDONA. Temperature is determined in each cell by evaluating the photon flux, identical to the ‘simple’ case used in ARTIS and TARDIS, but taking the gas and radiation temperatures to be equal, i.e. $T_e = T_R$. Ionisation and excitation are calculated assuming the Saha-Boltzmann equations, solved simultaneously for neutral to triply ionised ions for included elements between H to Zn, and doubly ionised onwards up to U (Tanaka & Hotokezaka 2013). The included elements in the compositions vary from study to study.

Works using this tool have focused largely on early time epochs relating to AT2017gfo (e.g. Tanaka et al. 2018; Domoto et al. 2021, 2022), or have limited themselves to LC analyses when extending to later times (e.g. Kawaguchi et al. 2021, 2022, 2023). While the RT in the TH13 code operates in 3D with time-dependent transfer, this comes at the expense of simplified photon-matter interactions relative to those modelled in SUMO. Much focus has been placed on the quality and calibration of atomic data in thee studies making use of the TH13 code, which has been used to identify key elements within the ejecta of
Figure 7.5: Comparison of synthetic spectra (blue) using the TH13 code to the spectra of AT2017gfo (black) at the 1.5, 2.5 and 3.5 day epochs. Potential features from key species are marked out. Taken from Domoto et al. (2022).

AT2017gfo, as shown in Figure 7.5.

7.5 NLTE Studies

Aside from the work presented in this thesis, very few studies have used full NLTE capable codes to model KNe out to later tines, when such modelling capacities are required. I present a brief overview of a few studies that have addressed NLTE effects to varying degrees of completeness and accuracy.

Hotokezaka et al. (2021) considered a relatively complete calculation of the NLTE gas state in a similar fashion to SUMO, with the additional assumption of an optically thin medium in order to simplify the radiation transfer, and for a single-zone ejecta model with a composition of pure neodymium (Nd). The temperature solution was found in the steady-state mode by equating the heating and cooling terms, where cooling was assumed...
Figure 7.6: Dielectronic and radiative recombination coefficients for Nd ions calculated using the atomic structure code HULLAC. A few auto-ionising states are found to dominate the recombination rates, typically significantly larger than the radiative recombination rates. Taken from Hotokezaka et al. (2021).

to be entirely by emission line cooling. The heating rate was obtained from radioactive energy deposition, following the simplified description of Axelrod (1980), instead of solving the full Spencer-Fano equation as in SUMO. There, 97% of the energy was attributed to heating, while the remaining 3% went towards NT ionisation. The collision strength treatment for allowed lines was similarly the Van Regemorter treatment, while a constant $\Upsilon = 1$ value was taken for forbidden transitions.

The ionisation structure was found by solving the steady-state rate equation, where ionisation came from $\beta$-decay electrons considering the 3% energy deposition going to NT ionisation, and recombination occurred from both radiative and dielectronic recombination. PI was partially taken into account by considering the probability that an emitted recombination photon will be reabsorbed and ionise a different or same species ion following the description in Axelrod (1980). Calculations of dielectronic and radiative recombination rates were conducted using the atomic structure code HULLAC (Bar-Shalom et al. 2001). There, it was found that this dielectronic recombination was the dominant process (See Figure 7.6) for Nd II - V.

Generally, this study is thus far the most similar to SUMO in the determination of ejecta conditions, and indeed the overall evolution of temperature and ionisation structure in the (steady-state) NLTE regime were found to be similar to those found in Paper I. We find the same temperature scaling for steady-state evolution, and also predict similar evolutions for ionisation structure, notably with respect to an expected ‘freeze-out’ once.
ionisation/recombination timescales become long. The actual values of these quantities differ, as different assumptions were made with respect to various microphysical processes e.g. recombination, ionisation, cross-sections of various processes etc., and their composition consisted solely of Nd. The choice of Nd is key, as it is expected to be one of the most complex lanthanide species, and contribute overwhelmingly to KN opacity (Tanaka et al. 2020). Due to its large number of line transitions, it is also expected to be an exceptional cooler, much more so than Te which dominated our model composition in Papers I and II. As such, some difference in solution is expected, but it is encouraging to note that the broad results of the NLTE studies conducted on KN so far appear to agree on the general evolution of quantities in the ejecta.

Other NLTE studies have used more semi-analytical approaches (Hotokezaka et al. 2022; Gillanders et al. 2023a; Hotokezaka et al. 2023), where thermodynamic properties such as temperature and ionisation structure are fixed. Excitation structure is however calculated with careful consideration of collision strengths for low lying levels of particular interest. This has been used to suggest the presence of Te III by emission of the forbidden 2.1 micron line, both in AT2017gfo (Hotokezaka et al. 2023) and the recent KN observation inferred from a sGRB afterglow excess (Levan et al. 2023), as shown in Figure 4.3 in Chapter 4.

Part of the difficulty of realistic NLTE modelling of KN is the lack of atomic data, not only for accurate energy levels and transitions, but also for the cross-sections of many (if not most) of the processes that must be modelled. While calibration of levels and transition wavelengths to what experimental data is available is an ongoing effort (e.g. Flörs et al. 2023), doing so for cross-sections will prove more difficult, in part due to the greater difficulties in measuring these quantities. As described in Chapter 6, SUMO currently uses mainly fitting formulae from various sources in order to approximately model the diverse processes included in the code. However, the accuracy of many of these processes is not well known, and is furthermore difficult to establish.

As an example, the Axelrod treatment for forbidden collision strengths has been found to give an underestimation of 2 orders of magnitude for the [Te III] 2.15 micron line described in Madonna et al. (2018); Hotokezaka et al. (2023). A recent study by Bromley et al. (2023) on the electron impact cross-sections of Pt I - III found that the Axelrod treatment systematically underestimates the collision strengths for forbidden lines, while the Van Regemorter treatment typically underestimates the allowed transition collision
strengths, though may also overestimate depending on electron temperature and ion. These preliminary results appear to imply that the formulae used in SUMO may in general be underestimated these collision strengths, though these studies are far from comprehensive, and it is ordinarily not possible to generalise diverse atomic properties to all atomic species.

It is possible that other processes such as PI, NT ionisation and excitation, and recombination may also be somewhat poorly estimated by the current treatments implemented in SUMO. However, given the lack of complete data sets of these cross-sections for r-process species, it is not yet possible to know the extent of these inaccuracies. In this sense, it is not necessarily guaranteed that NLTE studies of KNe will provide more accurate results than those making use of LTE methods and simplifications. However, making use of such approximations currently remains the best possible approach when attempting to study the NLTE regime of KNe on a broader basis than the consideration of a particular line from a given species. With time, the quality of theoretical atomic calculations, and availability of experimental data will both improve, and with those the accuracy of NLTE simulations. Current studies such as those presented in this thesis, and in particular Paper III, provide the first steps for fully consistent NLTE modelling, and in doing so, identify the key areas that need improvement, e.g recombination rates, electron impact collision strengths. This provides targets for the community on which to focus on first, providing impetus for further studies on these aspects.
8 Paper Summary

8.1 Paper I: On the Validity of Steady-State for Nebular Phase Kilonovae

Kilonovae evolve very rapidly compared to transients such as supernovae, and thus quickly reach the ‘nebular’ phase. This is characterised by NLTE conditions in the ejecta, and optically thin spectra dominated by atomic emission lines, as opposed to scattering lines over a quasi-blackbody continuum. This also corresponds to a drop in the photon diffusion time, such that the bolometric lightcurve in the early nebular phase tracks the instantaneous radioactive energy deposition.

Due to the dominance of atomic emission lines, the nebular phase offers a good opportunity for the identification of elements, or groups of elements in KN spectra, which may be more robust than abundances inferred from lightcurve and SED models alone. As such, correctly modelling the conditions inside the ejecta giving rise to the spectra, e.g. temperature, ionisation and excitation structure, is key to correctly inferring information about the KN and the elements therein.

This first paper tests a common assumption, that of steady-state, often used when modelling nebular phase SNe and KNe. Concretely, this approximation assumes that temperature, ionisation and excitation solutions in the ejecta are slowly changing, such that time derivatives of these quantities are set to zero. In other words, the non-thermal energy from radioactivity is rapidly reprocessed by thermal processes, and so the conditions at the current epoch have no knowledge of conditions at prior epochs.

While this is certainly the case for the early nebular phase, also called the steady-state phase, this will not be the case at later times, when thermal reprocessing becomes long. For temperature, cooling and heating eventually become inefficient, with associated
timescales becoming longer than the evolutionary time. For ionisation, these processes will be ionisation and recombination processes instead. Excitation is however expected to remain in steady-state due to the effect of fast, allowed radiative transitions. Should these time-dependent effects be important, it is expected that solutions of temperature and ionisation will deviate from those found by the steady-state equations.

The key question then is when this transition happens, and how significant it is within a reasonable observational time-frame. We try to answer this question by evolving a grid of KN ejecta models with varying mass and velocities, for a 4 element, solar abundance composition of tellurium (Te), cerium (Ce), platinum (Pt), and thorium (Th), up to 100 days after merger. We find that most parts of parameter space are insensitive to these effects, though find some onset of adiabatic cooling in the lowest density model. We also consider lower energy input by radioactive power, as well as inefficient thermalisation by lack of magnetic field trapping. These models show stronger effects, notably in the ionisation structure which is markedly less ionised. This leads to a cooler temperature, especially at later times when combined with adiabatic cooling. We also find a fainter bolometric LC by about 10% at 100 days.

These results imply that any late time modelling of KN may have to take such time-dependent effects into account if accurate inferral of ejecta properties is desired. Since much of the late time emission is expected to be in the infra-red from lanthanide-rich ejecta, and telescopes such as the JWST are able to observe KNe at late times, this paper provides a first insight as to what modellers must provide in order to fully understand current and future observations.

### 8.2 Paper II: NLTE Effects on Kilonova Expansion Opacities

A key component of nebular phase modelling for SNe and KNe is the use of NLTE physics when solving for temperature, ionisation and excitation. However, solving NLTE equations coupled with radiative transfer is computationally expensive, and requires the inclusion of many complex micro-physical processes. Many studies of KNe have thus far chosen to apply the LTE approximation to their calculations, which is essentially an application of the Saha-Boltzmann equation for a given temperature and density.
For LTE conditions to hold in the ejecta, the determination of the excitation structure of atomic species must be dominated by thermal collisional processes, over other processes such as spontaneous radiative deexcitation. This requires relatively high densities, which is expected to be the case in the ejecta right after merger, certainly up to 1-2 days after merger, roughly at the peak LC time. However, rapidly expanding media means that a detailed collisional balance may not be upheld after a few days, and thus NLTE conditions may be more applicable. This will be especially true for high lying states in large ions, which have many allowed transitions providing rapid deexcitation channels.

This second paper aims to establish how long the LTE approximation can be reliably applied to KN ejecta. We begin by verifying how accurately the Boltzmann equation manages to reproduce NLTE level populations, testing epochs from 3 to 20 days after merger. We find that LTE excitation for ‘standard’ ejecta parameters \( M_{ej} = 0.05 M_\odot, v_{ej} = 0.1c \), holds reasonably well up to about 5 days after merger, though high lying states in large ions are already depopulated by then in the NLTE solution. It has also been previously suggested that a strong thermal radiation field may help maintain LTE conditions for longer than thermal collisions (Kasen et al. 2017). This was also tested by considering level populations calculated in NLTE, but without radiative transfer. Concretely, this corresponds to having no photoionisations, radiative absorptions or stimulated emission, thus setting the terms in Equations 5.20, 5.32 and 5.33 to zero. Physically, this implies that the radiation field is sufficiently weak compared to collisional processes such that it is neglected, which we find to be reasonable at later epochs \( (t \gtrsim 5 \text{ days}) \). At early epochs however, we find that the radiation field plays a large role in pushing high lying states towards LTE, but this may sometimes come at the expense of lower lying states.

From excitation alone, it seems that LTE may be only applicable up to 5 days or so after merger. However, excitation structure is not easy to measure directly, but other quantities that depend on the excitation structure may be affected. Notably, expansion opacity, which is used extensively in LC and SED modelling is directly dependent on the excitation structure, and used to infer properties of the ejecta, such as component masses. As such, incorrect modelling of excitation structure may lead to inaccuracies in quantities derived from opacity.

We therefore calculate expansion opacities for three different excitation structure calculations: the Boltzmann equation (LTE), the full NLTE solution, and NLTE without radiative transfer in order to test the radiation field’s effect. We find that LTE expansion
opacities are reasonable again up to about 5 days, though tend to overestimate the true full NLTE values as time goes on. The NLTE without radiation field calculations yield expansion opacities practically identical to the full solution at all epochs, implying that these 'limited' NLTE opacities may perhaps be employed at late times.

We also test a full LTE model that makes use of the Saha ionisation equation. Since different ions have different contributions to opacity both in terms of wavelength and importance, a different ionisation structure leads to different opacity curves. In this case, we find that key ions missing from the Saha solution lead to drastic underestimations of opacity along key wavelength ranges. However, we also analytically show that the Saha equation is not expected to be valid for our models even at 3 days after merger, and so our results on this model should be taken as indicative, rather than definitive. Naturally, these results also depend on the ejecta type: lower density, more dynamic ejecta will fall out of LTE at very early times ($t \lesssim 3$ days), as the medium rapidly expands. The main conclusion remains however, that LTE excitation, and thus the early time modelling of AT2017gfo used by the community so far appears to be reasonable.

8.3 Paper III: NLTE Spectra of Kilonovae

Having tested the validity of the LTE and steady-state assumptions in the first two papers, Paper III focused on the generation of NLTE spectra of KNe. The aim of this study was not to identify particular features in the spectra of AT2017gfo, but rather identify which species play key roles in the spectral formation of KNe, and which processes drive this formation. In particular, the full fluorescence and resonance scattering treatment included in SUMO was useful to gauge the impact of these processes often omitted or simplified in other studies by the community.

To this end, three homogeneous composition models based on the nuclear networks of Wanajo et al. (2014) were created, with representative electron fractions of $Y_e \sim 0.35, 0.25, 0.15$. The energy deposition for each model was taken from the same nuclear network calculations, representing a notable upgrade in consistency compared to the first two studies, where energy deposition was parameterised. The ejecta models were identical, 5-zone, $\rho \sim v^{-4}$ set-ups, with a total ejecta mass of 0.05 $M_\odot$ and the velocity ranging from 0.05 – 0.3$c$ in fixed steps of 0.05$c$. These parameters were picked based on the hydrodynamical study of Kawaguchi et al. (2021), and represent a broadly 'average' ejecta
These models were run from 5 to 20 days after merger, with time-dependent mode activated from 10 days onwards based on the results of Paper I. In terms of thermodynamical results, the models follow the evolution generally found in Paper I. Temperature and ionisation fraction typically increase with time, with ejecta structure broadly stratified such that the outer layers are hotter and more ionised than the inner layers. The \( Y_e \sim 0.35 \) model, which naturally had lower energy deposition arising from the nuclear network calculations, exhibits strong time-dependent effects in the outer ejecta layers from 10 days onwards, corresponding to the low power/low density conditions identified in Paper I. Spectral effects are observable from 15 days onwards, showing that time-dependent effects may be relevant in observable times for fast moving, lanthanide-free ejecta components.

Considering the spectra themselves, it was generally found that line blocking from many optically thick lines is important below 1 micron for all models at early (5 days) times. The extent of the line blocking is stronger in elementally heavier compositions, i.e. the \( Y_e \sim 0.15 \) model had the most optically thick lines throughout the whole ejecta below 1 micron at 5 days. The emergent features at these early times are found in escape windows within the line blocking, rather than corresponding to intrinsically strong cooling lines. The line blocking decreases with time as the ejecta expand, leading to the emergence of new features at progressively bluer wavelengths.

The important extent of line blocking in the UV and optical means that blue photons cannot easily escape at these wavelengths, and are forced to fluoresce and scatter redwards. As such, it was found that fluorescence and scattering play key roles in KN spectral formation at least up to 20 days after merger. This also means that the ejecta were not completely optically thin up to 20 days, which is an important result for modelling NLTE epochs. Notably, assuming optically thin ejecta greatly simplifies the radiative transfer component of the modelling, and therefore it is useful to know that this assumption cannot yet be applied for the bulk ejecta in these epochs for the UV/optical/NIR wavelengths.

The critical role that fluorescence and scattering play during these epochs has important effects and implications. A major point is that the reprocessing of emission from these processes will change the SED shape in a non-thermal way. This implies that blackbody fits to the SED may have little physical meaning, as the gas and photon temperature are decoupled. This has implications for analyses that have fit blackbody continua to KNe spectra not only in the NLTE regime, but also potentially in the later LTE phase,
as fluorescence and scattering were also found to be key at earlier times (Shingles et al. 2023). Notably, gas temperatures estimated from BB fits to the radiation field may be underestimating temperature due to excess red flux from reprocessing.

As a result of line blocking and fluorescence/scattering, it was found that the model colours evolved from red to blue, contrary to the evolution of AT2017gfo, and most theorised colour evolutions of KNe. The explanation of this with respect to our models lies with consideration of the temperature and line blocking evolution. It has been well shown that temperature in the steady-state NLTE regime is expected to increase (Hotokezaka et al. 2021, and Paper I), which generally leads to bluer photons being emitted. However, at early times these photons are reprocessed redwards, and it is only at later times that they are able to escape in optical wavelengths. As such, these combined effects lead to models getting progressively bluer with time. Since the models in Paper III are homogeneous compositions, this further implies that real KNe do not have spherically symmetric, homogeneous ejecta, and that dimensional effects likely play a key role in providing escape paths for these bluer photons at early times (Shingles et al. 2023).

With respect to individual models, it was found that the first r-process peak elements Rb, Sr, Y and Zr dominate spectral formation in lanthanide-free ejecta ($Y_e \sim 0.35$ model). These elements on the left side of the periodic table are known to have strong allowed transitions due to their atomic structure, and the neutral and singly ionised species were previously identified in a photospheric context (Domoto et al. 2021). We also directly tested the impact of Sr II, the first robustly identified species in AT2017gfo (Watson et al. 2019; Domoto et al. 2021; Gillanders et al. 2022) by the 1 micron triplet transition, by removing it from the model. We found that the 1 micron triplet has significant impact on the local spectral shape, and minor effects from changes in the thermodynamic solutions were also observed at other wavelengths. From this model, we also found that the Rb I doublet from the first excited states to the ground state provides a very strong absorption channel that is optically thick throughout the whole ejecta up to 20 days. On the basis of this, we suggested that this may play a role in the formation of a P-Cygni feature seen in AT2017gfo around 760 nm, previously attributed to Y II lines (Sneppen & Watson 2023).

For the lanthanide-bearing ejecta ($Y_e \sim 0.25, 0.15$ models), we found that lanthanide species dominate spectral formation, in particular the elements Nd, Sm, and Dy. Identification of individual features here is difficult however, due to the complex nature of the lanthanides leading to relatively inaccurate atomic data. The importance of accurate data
is highlighted in the $Y_e \sim 0.15$ model, where we do not see the emergence of the [Te III] forbidden 2.15 micron line, claimed to be seen in AT2017gfo at 10 days (Hotokezaka et al. 2023), and in the long GRB afterglow excess at 29 days (Levan et al. 2023). There, our treatment of forbidden collision strengths from Axelrod (1980) yields a value two orders of magnitude smaller than the measured value (Madonna et al. 2018). This example highlights the need to incorporate measured atomic data where available, in order to provide accurate NLTE spectral modelling of KNe.

In general, this paper provided some of the first NLTE KN spectra that were fully modelled with complete NLTE machinery, and did not assume optical thinness in order to do so. Though the models and compositions remain somewhat simplified, this paper laid the foundation for further studies to develop along this line of research. Many follow up opportunities are available, from varying compositions, to inclusion of more accurate data pertaining to the various NLTE processes modelled. In effect, the first step towards fully consistent and accurate NLTE spectral modelling of KNe has now been taken.
The field of neutron star mergers and their associated emission began several decades ago, as a hypothetical site for the creation of r-process elements. Since then, much has evolved, especially in the past 6 years with the discovery of AT2017gfo. To this day, this event remains the only complete observation, and the only multi-messenger observation. The lack of combined gravitational wave and electromagnetic detections since then has not stopped the field from progressing, with theorists having ever more opportunities to explore new models and refine their techniques.

In this PhD thesis, the first two papers were dedicated to testing commonly used assumptions by the modelling community. The applicability of the LTE assumption was tested, and the misuse quantitatively assessed with respect to the broadly employed expansion opacity formalism. The consideration of steady-state for the NLTE regime was also tested, which necessitated the development of new code in SUMO, which is now (at least thus far, to the author’s knowledge), the only code to consider time-dependent effects in the determination of the ejecta gas state. The final paper included in this thesis generated the first fully consistent NLTE spectra of KNe, albeit from rather simplistic toy models. Even so, this allowed the determination of the species that play key roles in KN spectral formation, as well as the identification of resonance scattering and fluorescence as vital transfer processes. Furthermore, it was determined that the KN ejecta may not be optically thin immediately after exiting the LTE regime, and that KNe may not transition into the optically thin nebular phase directly after the LTE photospheric phase.

As explored in Chapter 7, various radiative transfer simulations have been used in conjunction with hydrodynamical and nuclear network modelling in order to provide KN models. These have been generated with various assumptions, simplifications and differing treatments of microphysics. However, one theme remains constant throughout all the radiative transfer modelling efforts: we need to know more about the r-process elements.
themselves. Atomic energy levels and transitions are used by all models, whether in LTE or NLTE, and complete data sets are currently taken from theoretical calculations, as experimental measurements for many r-process species are incomplete or even absent. This lack of atomic data is even worse when considering the cross-sections needed to model the various microphysical processes that play key roles in NLTE modelling.

In order to improve the accuracy of NLTE modelling for future studies, the radiative transfer community must work hand in hand with the atomic physics community, to establish what is needed and to what precision. Fortunately, such undertakings have already begun in earnest, offering a promising future for KN modelling. With models therefore improving every day, the author now hopes that some day, they will be applied to new, full observations of neutron star mergers and kilonovae.


Acknowledgements

At the surface once emerged
Radiant in birth
Free at last, in its final form
The photon dashed for Earth
- Be’lakor, An Ember’s Arc

I would like to thank my supervisors Anders Jerkstrand and Stephan Rosswog for their guidance and aid through the PhD and these past four years. I in particular want to thank Anders for taking a chance by recruiting me in the beginning when I had no experience in transients or radiative transfer modelling whatsoever. I want to thank my mentor Jaime de la Cruz Rodriguez, who was always available when I needed to talk to him, and always proved a capable and attentive mentor. I am grateful to my thesis committee, who made the time to come to Stockholm on a late December 18th right before Christmas.

I also want to thank Jeremy Drake and Cecilia Garraffo, for giving me the opportunity to enter the world of astrophysical research right after my bachelor’s degree. The summer spent doing research at the Harvard-Smithsonian CfA crystallised my desire to make this my career. To my master’s supervisor Kevin Schawinski, thank you for accepting me into your group at the time, and making me realise my passion for compact objects and high energy astrophysics. To my master’s mentor Benny Trakhtenbrot, thank you for being present and supporting me through the master’s thesis until the publication of our study, despite the tough environment at the time.

I am grateful for the friends I made in the department, and all the people I’ve had the chance to meet and interact with. To Ivo, thank you for being the best office mate and best office ‘husband’ anyone could wish for, I look forward to you becoming recognised as a respected authority in your field soon. To Moa, all the best with the rest of your
PhD, and being the simultaneous face of Swedish astronomy and martial arts, happy to say I know a real celebrity. To Cis, for being an endless source of silly fun, and a friend with whom I could share my music taste. To Gloria, for being an excellent vice-president and friend, please don’t forget me when you become a famous professor! To Alex and Giacomo, thank you both for hosting so many parties, dinners and hangouts, and for making Norwegian Reggaeton perpetually stuck in my head. To Aurore, who preceded me in her defense and start of the post-doc, I hope we will once again be able to have our late night hang outs over a couple beers. To my younger PhD brothers Bart and Stan, I hope your PhDs go well, thank you for the technical talks, and Stan for translating things to Swedish for me. To Jack, thank you for being one of the few physics friends with whom I could also share my passion for volleyball, hope you keep training (responsibly)! To the rest of the physics people I’ve had the pleasure of knowing, Mila, Gayathri, Mattia, Alice, Suhyeon, Aaron, Oliver, Jenny, Nikhil, Conor, Alex (Le Reste), I wish you all the best for the future, and hope to see you again one day.

To my friends outside of physics, in particular my volleyball circle, I want to thank you all for the good times we’ve had. I never thought that I would one day end up on a competitive sports team playing in a real, official division, so thank you to the Elden team for having me be a part of that experience. Cheers to our coach Amna, without whom the team would never exist. To the players and my comrades on the court: Stefanos, Mo, Axel, Pyry, Carl, Allan, Lucas, Peter, Mando, Patrik, Thomas and Cyrus, may you enjoy many more victories! To Don, I’m sure our travels will overlap again at some point, so here’s to our next meeting somewhere on the globe. To Fernando, thank you for being a great guy, I’ll never forget our first meeting when you broke my bottle with your serve! To Kiril, my first volleyball friend, thank you for remaining a great friend this entire time, both on and off the court. To Julija, my wacky dancing friend, never forget that waving your arms like wet noodles is a perfectly good dance move, I hope you will keep representing it on the Swedish dance floors. To Natalia, sportswoman extraordinaire, thank you for being the life of conversation so many times, what would we do without you? To Amir, thank you for organising our Saturday volleyball sessions for so long, mister tiger! To Andreas, enjoy being able to eat and drink all you want while you can, and let me know when you are finally unable to eat three pizzas one sitting. To Vedrana, thank you for being another person on the court who shared my reactions and thoughts so many times, it was nice to feel understood without having to say a word.
Αγαπητή Άννα, Λεμονάκι μου. Όπως με ρώτησες πολλές φορές, ποια θα ήταν η ζωή μου χωρίς εσένα; Βαρετό! Σε ευχαριστώ λοιπόν ότι ήσουν η καλύτερη φίλη που θα μπορούσα ποτέ να έχω, με στηρίζει στις πιο δύσκολες στιγμές και γιορτάζει μαζί μου στις καλύτερες στιγμές. Πολλά φιλιά και πολλή αγάπη για σένα, αγάπη μου.

I also want to thank my old friend Matthias, with whom I had many talks and calls over these past four years. Ever since school you’ve been a great friend, and will always be the "good man". Can’t wait to meet up again sometime soon buddy!

To my sister Diane, who in some ways could not be more different to me, and yet so similar in others. Thank you for all the good moments growing up, I hope you reach all your goals. When you make a red carpet premier, please don’t forget to invite me!

A mes parents, je vous remercie de m’avoir soutenus tout au long du trajet, pas toujours évident. C’est grâce à vous, à votre support et amour que je suis qui je suis aujourd’hui. J’ai hâte de partager la prochaine étape du voyage avec vous, et pour bien encore des années après.
Bibliography

Cowan, J. J., Sneden, C., Lawler, J. E., et al. 2021, Reviews of Modern Physics, 93, 015002
Cowan, R. D. 1981, Phys. Scr, 24, 615
Eddington, A. S. 1925, The Observatory, 48, 73
Gu, M. F. 2008, Canadian Journal of Physics, 86, 675
Hawking, S. W. & Israel, W. 1989, Three Hundred Years of Gravitation
Hotokezaka, K., Piran, T., & Paul, M. 2015, Nature Physics, 11, 1042
Huba, J. 2013, Naval Research Laboratory, Washington, DC
Jerkstrand, A. 2011, PhD thesis, Department of Astronomy, Stockholm University
Jönsson, P., Gaigalas, G., Bieroni, J., Fischer, C. F., & Grant, I. 2013, Computer Physics Communications, 184, 2197
Kasen, D., Metzger, B., Barnes, J., Quataert, E., & Ramirez-Ruiz, E. 2017, Nature, 551, 80
Oertel, M., Hempel, M., Klähn, T., & Typel, S. 2017, Reviews of Modern Physics, 89, 015007
Osterbrock, D. E. & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei
Perego, A., Thielemann, F. K., & Cescutti, G. 2021, in Handbook of Gravitational Wave Astronomy, 1
Postnov, K. A. & Yungelson, L. R. 2006, Living Reviews in Relativity, 9, 6
recipes in FORTRAN. The art of scientific computing
Radice, D., Bernuzzi, S., & Perego, A. 2020, Annual Review of Nuclear and Particle
Science, 70, 95
Rutten, R. J. 2003, Radiative Transfer in Stellar Atmospheres
Ryabchikova, T., Piskunov, N., Kurucz, R. L., et al. 2015, Physica Scripta, 90, 054005,
Rybicki, G. B. & Lightman, A. P. 1979, Radiative processes in astrophysics
Schmidt, P. 2020, Frontiers in Astronomy and Space Sciences, 7, 28
Shibata, M. & Hotokezaka, K. 2019, Annual Review of Nuclear and Particle Science, 69, 41
Sobolev, V. V. 1960, Moving envelopes of stars
Spencer, L. V. & Fano, U. 1954, Physical Review, 93, 1172