

Electronic Thesis and Dissertation Repository

6-15-2021 10:30 AM

Mathematical Modelling & Simulation of Large and Small Scale Structures in Star Formation

Gianfranco Bino, *The University of Western Ontario*

Supervisor: Shantanu Basu, *The University of Western Ontario*

Co-Supervisor: David Jeffrey, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Applied Mathematics

© Gianfranco Bino 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Numerical Analysis and Computation Commons](#), [Other Applied Mathematics Commons](#), and the [Stars, Interstellar Medium and the Galaxy Commons](#)

Recommended Citation

Bino, Gianfranco, "Mathematical Modelling & Simulation of Large and Small Scale Structures in Star Formation" (2021). *Electronic Thesis and Dissertation Repository*. 7860.
<https://ir.lib.uwo.ca/etd/7860>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

This thesis aims to study the magnetic and evolutionary properties of stellar objects from the prestellar phase up to and including the late protostellar phase. Many of the properties governing star formation are linked to the core's physical properties and the magnetic field highly dictates much of the core's stability.

The thesis begins with the implementation of a fully analytic magnetic field model used to study the magnetic properties governing the prestellar core FeSt 1-457. The model is a direct result of Maxwell's equations and yields a central-to-surface magnetic field ratio in the equatorial plane in cylindrical coordinates. The model additionally gives estimates to the field direction and the relative magnitude at all points within the core. Through non-linear fitting, the single plane magnetic field is fit to observationally measured near-infrared polarization segments resulting from the integrated properties of the magnetic field and dust grains. The model is back tested by deploying the radiative transfer code POLARIS, used to simulate synthetic polarization maps resulting from the integrated scattering and emission properties of dust grains.

This study is further extended to the protostellar phase in the core's evolution. In order to accurately model the dense, star forming regions of protostellar cores, non-ideal magnetohydrodynamic (MHD) simulations are used by solving the resistive MHD equations on a nested grid at a series of length scales. The simulation is used to study three instances in time for the core's evolution represented in terms of the core's mass. This methodology enables one to resolve the central region on small scales and is more computationally efficient than the adaptive grid method. In doing such, the simulation outputs are served as inputs to the POLARIS code used to simulate synthetic polarization maps at all length scales within the protostellar core allowing for robust, high resolution outputs of the polarization segments within the small scale regions of the core. Additionally, the synthetic polarization maps are used to model the polarization state of the Orion Source I cloud. Through superimposition and an analysis on the mean field direction, the emission or absorption properties of the observational measurements from ALMA can be inferred from both the POLARIS and MHD simulations.

The late protostellar evolution phase involves studying the dynamics of the episodic behavior of mass accretion. By implementing an optimized form of recurrent neural networks utilizing the reservoir computing framework, one can produce robust, model-free predictions to uncover and extract important, underlying temporal features governing mass accretion in the late evolution period. Through carefully pre-processing data, the network is run on a series of hydrodynamic and stellar evolution simulations in an attempt to successfully recover the expected behavior of the simulations whilst introducing an effective method for performing predictive analysis on data that has no governing model.

Summary for lay audience

Stars have long been known to be an important factor in the evolution of the universe as a whole, including in its constituent parts like galaxies and planets. In this thesis, the formation of stars is studied in different evolutionary epochs from the prestellar phase up to and including the formation and evolution of protostars. A critical component in star formation is the magnetic field. A large part of this thesis aims to study the magnetic field that governs both prestellar and protostellar cores. Using theoretical simulations in conjunction with observational data, robust models are developed that give insight to some important stellar properties governing the cores. This thesis additionally provides an introduction to forecasting time series using a new form of neural networks.

Acknowledgements

I would like to first and foremost extend my sincere gratitude to my advisor, Dr. Shantanu Basu. His continued guidance, knowledge and support is something that I was fortunate to have throughout my graduate studies. I learnt more than I thought I could and he has played a large role in getting me to where I am today.

I would also like to thank my collaborators Dr. Masahiro Machida and Dr. Lyle Muller for their consistent involvement and insight. I've had the opportunity to learn from them and their expertise have contributed to a large part of my work.

I am also thankful for the staff and all of the Applied Mathematics department at the University of Western Ontario for providing continuous resources, support, the ability to work with incredible professors, students and for giving me a phenomenal experience.

I would like to also express appreciation and love to my entire family for their continuous support and belief in me. My mother, Lucy Del Corpo and father, Vince Bino have been a everything anyone can ask for in parents and I will be forever grateful for them.

Contents

Abstract	i
Summary for lay audience	ii
Acknowledgements	iii
Introduction	1
1 Overview	1
2 Prestellar Cores	2
2.1 Hourglass Magnetic field	2
2.2 Magnetohydrodynamics	3
2.3 Gravitational contraction	4
2.4 Flux freezing	5
3 Protostellar Cores	7
3.1 Early Protostellar Formation	7
3.2 Mass Accretion & Episodic behavior	7
4 Chapter Preview	9
Chapter II : Large Scale Structure & Prestellar cores	11
5 Introduction	12
6 Mathematical Model	13
6.1 Model Description	13
6.2 Parameters & Reformulation	14
7 Polarimetry & Model Fit	17
7.1 Polarization Vectors	17
7.2 Polarimetry of FeSt 1-457	19
8 POLARIS Pipeline	27
8.1 Stokes Vector and Polarization	27
8.2 Radiative Transfer Equation	31
8.3 Monte Carlo Photon Transfer	32
8.4 Grain Alignment	34
8.5 Intensity and Polarization Maps	36

9	RT Simulation Model	38
9.1	Grid Properties	38
9.2	Physical Properties	39
9.3	Density Profile	40
10	RT Simulation Output	42
10.1	Dust temperature distribution	42
10.2	Dust thermal emission with grain alignment	43
10.3	Mass-to-Flux ratio	44
11	Incorporating Radial Current Densities	50
12	Conclusion	52
 Chapter III : Small Scale Structure & Protostellar cores		53
13	Introduction	54
14	Non-Ideal Magnetohydrodynamic Simulation	55
14.1	Initial State	55
14.2	Ohmic Dissipation & Sink Treatment	55
14.3	Mathematical Methods	56
14.4	MHD Simulation Output	58
15	Numerical MHD in POLARIS	71
15.1	Cartesian Grid with Refinement	71
15.2	Model Properties	72
16	Simulation Results	73
16.1	Synthetic Polarization Maps	73
16.2	Modelling the Polarization of Orion Source I	77
17	Conclusion	84
 Chapter IV : Episodic Accretion & Mass Accumulation		85
18	Introduction	86
19	Hydrodynamics & Stellar Evolution	87
19.1	Initial State	88
20	Modelling Chaotic Time Series	89
20.1	Chaos & Maximum Lyapunov Exponent	89
20.2	Neural Networks	90
20.3	Reservoir Computing & Echo State Neural Networks	93

20.4	Data Pre-processing	95
20.5	Network Architecture	96
20.6	Hyperparameter Optimization	98
20.7	Predictive Analysis	100
20.8	Prediction Bands	104
20.9	Robustness & Hybrid Networks	105
21	Applications, Conclusion & Next Steps	110
21.1	FUor & EXor Objects	110
21.2	Caveats & Remedies	110
21.3	Next Steps	111
	Conclusion	113
	Bibliography	115
	Appendix	123
A1.1	Magnetic Field Model Mathematical Derivation	123
A1.2	Dimensional Analysis for magnetic field solution	132
A1.3	SQP Algorithm	133
A1.4	MHD Gravitational Solver	136
A1.5	Eckmann Algorithm	139
A1.6	OWA-ESN Outputs for Models 26 to 32	141

List of Figures

1	Hourglass magnetic field.	2
2	Flux Freezing demonstration.	6
3	Schematic figure of the evolution of a star-forming core.	7
4	Basu and Ewertowski replicated Hourglass Plot.	14
5	Background field $B = 1.5B_0$.	15
6	Background field $B = 0.5B_0$	15
7	Background field $B = 0.25B_0$	15
8	Schematic figure of the polarization properties of light.	17
9	Near-infrared Polarimetry of FeSt 1-457.	18
10	Field Density and Streamlines with SQP fit parameters (Full plane fit).	19
11	B_r/B_0 and B_z/B_0 model plots (Full plane).	20
12	Field Density and Streamlines with SQP fit parameters (Left plane fit).	21
13	B_r/B_0 and B_z/B_0 model plots (Left plane).	22
14	Field Density and Streamlines with SQP fit parameters (Right plane fit).	23
15	B_r/B_0 and B_z/B_0 model plots (Right plane).	24
16	Field Density and Streamlines with SQP fit parameters (Fixing $R = 0.2$ pc).	25
17	B_r/B_0 and B_z/B_0 model plots (Fixing $R = 0.2$ pc).	25
18	Plane of sky Electric field vector.	27
19	Plane of sky Electric field vector in three coordinates.	28
20	Schematic figure of the polarization ellipse.	28
21	Poincaré sphere spanning Stokes parameters.	30
22	Grain's disrupted alignment with the magnetic field.	34
23	Cylindrical grid generation.	38
24	Polarization map comparisons of the full plane fit and $R = 0.2$ pc fitted model.	44
25	Polarization maps for the left and right plane models.	45
26	Dust grain comparison with polarimetry data.	46
27	Statistical figures assessing the model's goodness of fit.	49
28	Magnetic field morphology as a function of scale length ratio.	51
29	Exemplary geometry of a nested grid demonstrated for a total of 5 levels of refinement.	57
30	MHD Simulation - Time evolution of the outflow region in terms of density.	58
31	MHD Simulation - Temperature for slice in x , y and z plane for $M = 0.01M_\odot$.	59
32	MHD Simulation - Density for slice in x , y and z plane for $M = 0.01M_\odot$.	60
33	MHD Simulation - Magnetic Field - $M = 0.01M_\odot$.	61
34	MHD Simulation - Velocity Field - $M = 0.01M_\odot$.	62
35	MHD Simulation - Temperature for slice in x , y and z plane for $M = 0.2M_\odot$.	63
36	MHD Simulation - Density for slice in x , y and z plane for $M = 0.2M_\odot$.	64
37	MHD Simulation - Magnetic Field - $M = 0.2M_\odot$.	65

38	MHD Simulation - Velocity Field - $M = 0.2M_{\odot}$.	66
39	MHD Simulation - Temperature for slice in x , y and z plane for $M = 0.28M_{\odot}$.	67
40	MHD Simulation - Density for slice in x , y and z plane for $M = 0.28M_{\odot}$.	68
41	MHD Simulation - Magnetic Field - $M = 0.28M_{\odot}$.	69
42	MHD Simulation - Velocity Field - $M = 0.28M_{\odot}$.	70
43	Exemplary tree diagram of an adaptive octree grid.	71
44	Synthetic polarization maps for $M = 0.01M_{\odot}$.	74
45	Synthetic polarization maps for $M = 0.2M_{\odot}$.	75
46	Synthetic polarization maps for $M = 0.28M_{\odot}$.	76
47	Polarization map for Orion Source I.	78
48	Synthetic polarization map for the best representative model.	79
49	Synthetic polarization maps for $M = 0.01M_{\odot}$ with ALMA data.	81
50	Synthetic polarization maps for $M = 0.2M_{\odot}$ with ALMA data.	82
51	Synthetic polarization maps for $M = 0.28M_{\odot}$ with ALMA data.	83
52	Mass accretion rate versus time.	87
53	Schematic figure of a one layer perceptron.	90
54	Schematic figure of a multilayered neural network.	91
55	3-layer recurrent neural network.	93
56	Schematic figure for the structure of an echo state neural network.	95
57	Schematic figure for the architecture of the optimal weighted average ESN.	98
58	The training/validation/testing segments for model 29.	101
59	Optimized output for $\dot{\mathcal{M}}(t)$ and $\dot{M}_d(t)$.	102
60	Optimized output for the final output $\dot{M}(t)$.	102
61	The time evolving RMSE up 46.1 Lyapunov times.	103
62	The RMSE demonstrated as a function of the number of reservoirs in the series.	103
63	The distribution of outputs from the reservoir series.	104
64	The OWA-ESN output for $\dot{\mathcal{M}}(t)$ and $\dot{M}_d(t)$ with prediction bands.	106
65	The OWA-ESN final output for $\dot{M}(t)$ with prediction bands	106
66	Schematic figure of the hybrid network.	107
67	Out of sample outputs for the network's deterministic and fluctuating components.	109
68	Out of sample free run for the network's final output.	109
69	Simulation FMG scheme.	138
70	Optimized output for model 26.	142
71	Optimized output model 27.	143
72	Optimized output for model 28.	144
73	Optimized output for model 30.	145
74	Optimized output for model 31.	146
75	Optimized output for model 32.	147

List of Tables

9.1	Physical and Numerical Model Properties.	39
9.2	Additional Properties.	41
14.1	Nested Grid Properties.	57
15.1	Protostellar Properties.	72
16.1	Table representative of the respective mean offset angles for each mass instance.	80
19.1	Initial parameters for models 26-32.	88
20.1	Maximum Lyapunov exponent for each respective model.	89
20.2	Search Space for the grid search optimization.	98
20.3	The optimal hyperparameters for all models.	100
20.4	Summary of the respective performance metrics for model 29.	101
20.5	Hybrid network hyperparameters.	105
A1.1	Summary of the OWA-ESN performance metrics for models 26 to 32.	141

INTRODUCTION

1 Overview

Stars are known to play a key role in many cosmological processes ranging from the evolution of the early universe to galaxy evolution and the formation of planets [1, 2]. Therefore, understanding star formation becomes an integral tool in understanding the universe as a whole. Many properties surrounding star formation can be inferred by studying molecular clouds. Molecular clouds are highly dense interstellar clouds made up of predominantly molecular hydrogen. Clumps within the cloud combined with sufficient gravity provide the beginning stages of star formation. Early in the stellar evolution timeline, prestellar cores form as a hydrostatic starless region. As the core accrues mass from the surrounding cloud, the core's self-gravity begins to dominate the internal pressure causing a collapse which leads to the formation of a protostar and a protostellar core. Magnetic fields play a large role in the formation and stability of the core [3, 4]. Understanding the magnetic properties of pre and protostellar cores can lead to understanding some of the stellar properties such as the contraction mechanisms governing the collapse. This thesis will provide a comprehensive study on the magnetic and dynamical properties governing stellar evolution from prestellar formation up to and including late protostellar evolution.

2 Prestellar Cores

2.1 Hourglass Magnetic field

The hourglass magnetic field results from gravitational contraction of plasma that is governed by the flux freezing property within the core enabling the field lines to be pulled toward the contracting region and gives rise to the morphology in figure 1. An hourglass pattern may also arise from magnetic induction in a contracting cloud in which non-ideal MHD processes are at work. Basu et al. [5] showed that the degree of curvature of an hourglass pattern would depend on the amount of non-ideal MHD effects at work. One approach to fitting an hourglass morphology is with a set of nested parabolas, which is a purely mathematical matching. Numerical MHD models yield more physically realistic hourglass shapes but are not available in a closed analytic form. Hourglass morphology has been observed in NGC 1333 IRAS 4A [6] and G31.41+0.31 [7] from dust polarization measurements. Polarimetry is a powerful tool to elucidate magnetic field morphology and reveal hourglass or other systematic magnetic field patterns.

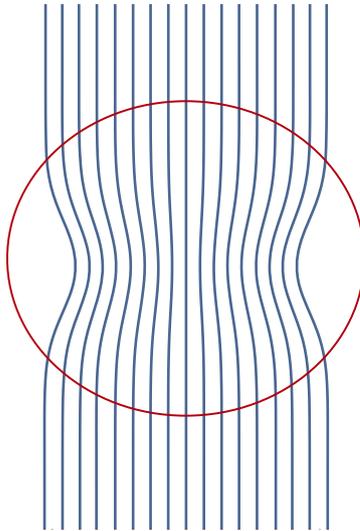


Figure 1: Standard structure of the field lines governing an hourglass magnetic field.

This has been known since Davis Jr and Greenstein [8] proposed that elongated paramagnetic dust grains would generally spin about their minor axis and that this axis would come into alignment with the ambient magnetic field through paramagnetic relaxation (grain alignment theories are discussed further in section 8.4 of chapter 2). Although direct magnetic field measurement is difficult, various properties of the magnetic field can be determined by the morphology.

2.2 Magnetohydrodynamics

Magnetohydrodynamics (MHD) models the magnetic properties of conducting fluids and plasmas. They couple the dynamics of Navier-Stokes equations with the electromagnetic properties of Maxwell's equations. Magnetic fields can induce currents within plasmas which in turn generates its own magnetic field giving a feedback effect in changing the inducing field. Astrophysical plasmas exist in prestellar cores which drives contraction and pinches the field lines within the core which produces the hourglass pattern [9]. The full, non-ideal MHD equations are given as:

$$(2.1) \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

$$(2.2) \quad \rho \frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla p + \frac{\mathbf{j} \times \mathbf{B}}{c} + \rho \mathbf{g} + \rho \nu \nabla^2 \mathbf{v}$$

$$(2.3) \quad \frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B} + \nabla \times (\eta_A [(\nabla \times \mathbf{B}) \times \mathbf{B}])$$

$$(2.4) \quad \frac{\rho^\gamma}{\gamma - 1} \frac{D}{Dt} \left(\frac{p}{\rho^\gamma} \right) = \nabla \cdot (\kappa \nabla T) - \rho^2 Q(T) + \frac{j^2}{\sigma} + H$$

which are representative of mass continuity, momentum conservation, magnetic induction and temperature evolution respectively. Here, $D/Dt \equiv (\partial_t + \mathbf{v} \cdot \nabla)$ is taken as the material derivative and there is the additional no monopole constraint that $\nabla \cdot \mathbf{B} = 0$. Plasmas in prestellar cores are highly conductive [10] and so the resistive term, η in equation 2.3 becomes:

$$\eta = \frac{c^2}{4\pi\sigma} \approx 0 \quad \text{for } \sigma \rightarrow \infty$$

Additionally, due to the property of flux freezing assumed in ideal MHD, the magnetic field is fully coupled to the neutral material, ignoring the motion of charged particles. Therefore, the coefficient of ambipolar diffusion $\eta_A = 0$ and so equation 2.3 becomes:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B})$$

The plasma is therefore said to have a high magnetic Reynolds number which is the ratio of induction to diffusion processes. This suggests that the fluid is highly inductive and therefore diffusive contributions from the viscosity ν can also be ignored. Ideal MHD assumes no additional heating which makes the system adiabatic and so all the temperature dependances in equation 2.4 can be ignored (including H which represents the sum of all other heating sources) and so equation 2.4 becomes (recalling the high conductivity limit has $j^2/\sigma \rightarrow 0$):

$$\frac{\rho^\gamma}{\gamma - 1} \frac{D}{Dt} \left(\frac{p}{\rho^\gamma} \right) = 0$$

and so the ideal MHD equations are given as:

$$(2.5) \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

$$(2.6) \quad \rho \frac{\partial \mathbf{v}}{\partial t} + \rho (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla p + \frac{\mathbf{j} \times \mathbf{B}}{c} + \rho \mathbf{g}$$

$$(2.7) \quad \frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B})$$

$$(2.8) \quad \frac{\rho^\gamma}{\gamma - 1} \frac{D}{Dt} \left(\frac{p}{\rho^\gamma} \right) = 0$$

Prestellar core plasmas share the same properties discussed in the ideal MHD limit and thus become a natural way to model the magnetic field in these systems. These properties lead to gravitational contraction within the core and the flux freezing property of the fluid which lead to observed hourglass patterns in NGC 1333 IRAS 4A [6] and G31.41+0.31 [7].

2.3 Gravitational contraction

Gravitational contraction plays an important role in the formation of hourglass patterns. This occurs primarily when the magnetic pressure is insufficient to balance out gravity [11]. The magnetic pressure is given from equation 2.6 on the right hand side. Taking:

$$(2.9) \quad \mathbf{j} \times \mathbf{B} = \frac{c}{4\pi} (\nabla \times \mathbf{B}) \times \mathbf{B}$$

and making use of the vector identity:

$$\nabla(\mathbf{X} \cdot \mathbf{Y}) = (\mathbf{X} \cdot \nabla) \mathbf{Y} + (\mathbf{Y} \cdot \nabla) \mathbf{X} + \mathbf{X} \times (\nabla \times \mathbf{Y}) + \mathbf{Y} \times (\nabla \times \mathbf{X})$$

the Lorentz force can be described in two parts:

$$(2.10) \quad \underbrace{\mathbf{j} \times \mathbf{B}}_{\text{Force}} = \underbrace{\frac{c}{4\pi} (\mathbf{B} \cdot \nabla) \mathbf{B}}_{\text{Tension}} - \underbrace{\nabla \frac{B^2}{8\pi}}_{\text{Pressure}}$$

The pressure is driven in the direction of weaker field strength and the tension is directed radially inward to field line curvature. Gravitational contraction occurs when the core is super critical. That means the mass to flux ratio exceeds the critical value of:

$$(2.11) \quad \left(\frac{M}{\Phi} \right)_{\text{crit}} = \frac{1}{2\pi\sqrt{G}}$$

A super critical core would suggest M to be the dominant contributor and so when the magnetic flux is the weaker of the two; the contribution from the magnetic pressure is being dominated by the force of gravity. This causes gravity to 'pinch' the field lines toward the center of the core. The field lines at this point remain distorted in the hourglass shape due to the magnetic tension strengthening. The tension will also point opposite to gravity and so eventually will balance the gravitational contraction and provide support to uphold the hourglass morphology.

2.4 Flux freezing

An important property for ideal MHD systems is the notion that the fluid is said to be tied to the magnetic field lines. This causes the fluid-field topology to be fixed. The principle behind this comes from equation 2.7 and relating it to Faraday's law. Taking the flux of the magnetic field through some surface S to be:

$$\Phi = \int_S \mathbf{B} \cdot d\mathbf{S}$$

The flux of the magnetic field comes from two contributors; the change in \mathbf{B} over the surface S and the motion of the plasma element contour C . The first contribution can be given as:

$$\left(\frac{d\Phi}{dt}\right)_1 = \frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{S} = \int_S \frac{d\mathbf{B}}{dt} \cdot d\mathbf{S}$$

$$\left(\frac{d\Phi}{dt}\right)_1 = -c \int_S \nabla \times \mathbf{E} \cdot d\mathbf{S}$$

where the Maxwell-Faraday's law has been applied to introduce the electric field component in cgs units. The second contribution states that if the contour C is moving with some element $d\mathbf{l}$, then the contour will sweep an area $d\mathbf{S} = \mathbf{v} \times d\mathbf{l}$ per unit of time. Thus the second contribution is given as:

$$\left(\frac{d\Phi}{dt}\right)_2 = - \int_C \mathbf{B} \cdot \mathbf{v} \times d\mathbf{l} = - \int_C \mathbf{B} \times \mathbf{v} \cdot d\mathbf{l}$$

applying Stokes theorem

$$\left(\frac{d\Phi}{dt}\right)_2 = \int_S \nabla \times (\mathbf{B} \times \mathbf{v}) \cdot d\mathbf{S}$$

and so:

$$(2.12) \quad \left(\frac{d\Phi}{dt}\right) = \left(\frac{d\Phi}{dt}\right)_1 + \left(\frac{d\Phi}{dt}\right)_2 = - \int_S \nabla \times (c\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot d\mathbf{S}$$

$$(2.13) \quad = -c \int_S \nabla \times \left(\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c}\right) \cdot d\mathbf{S}$$

The Ohmic MHD law is given as:

$$(2.14) \quad \mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c} = \eta \mathbf{j}$$

where for ideal conditions, $\eta = 0$ and so:

$$\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c} = 0$$

therefore equation 2.13 becomes:

$$\left(\frac{d\Phi}{dt}\right) = 0$$

which shows that the flux of field lines through the plasma due to the motion of plasma and change in magnetic field remains unchanged. So one can imagine having a region S that is threaded with a uniform magnetic field as shown in the left diagram in figure 2.

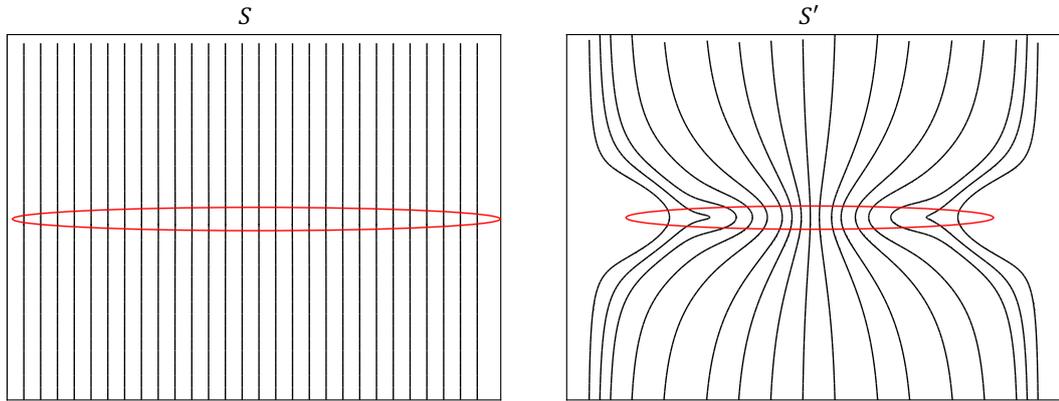


Figure 2: Left: Field lines through S are uniform **Right:** Field lines through S' form an hourglass due to flux freezing.

If flux freezing applies and the region transforms from $S \rightarrow S'$, the flux of field lines threading the region will be conserved and so can change the magnetic field morphology to that of an hourglass pattern as shown in the right diagram in figure 2. The consequences of gravitational contraction and flux freezing in prestellar plasmas form the ideal recipe for an environment that is prone to forming hourglass patterns.

3 Protostellar Cores

3.1 Early Protostellar Formation

As the core evolves, collapse occurs and forms a protostellar core. A protostellar core is the evolutionary stage following prestellar cores. It contains a young embryonic star that is still gathering mass from its parent gas cloud. This takes the core's evolution from the dense core phase into the earliest stages of star formation. The general idea of this transition is shown in figure 3 [12].

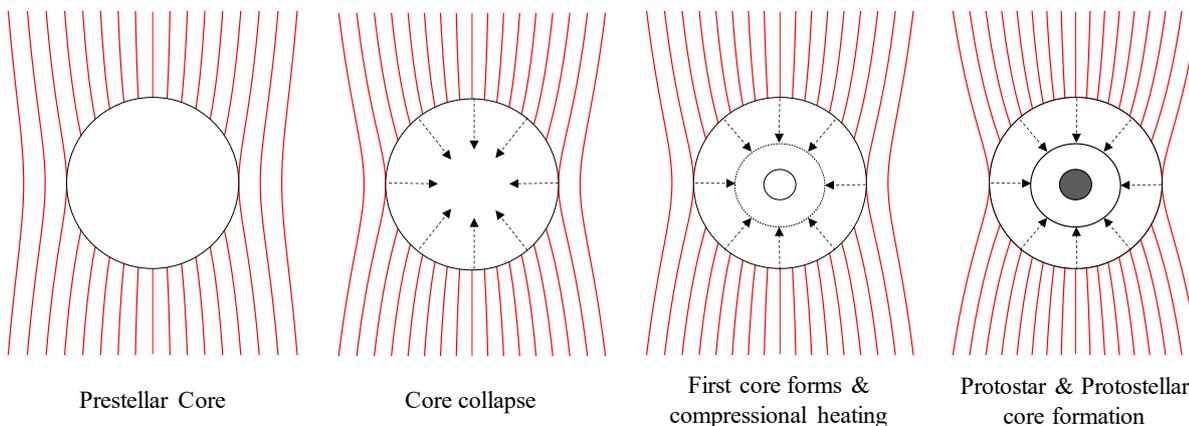


Figure 3: Schematic figure of the evolution of a star-forming core.

The initial collapse comes from the core exceeding some mass threshold $M_{\text{core}} \geq M^*$ where M^* is the critical mass at which collapse starts to take place. The core accretes mass from the surrounding gas and becomes sufficiently dense as to collapse and form a central object. The formation of a central object introduces some important differences to modelling protostellar cores. Primarily, the clouds become lightly ionized [13] and the magnetic field can no longer be coupled to the neutral material. This means that the flux frozen property used in modelling prestellar cores is no longer valid and one must revert to some form of non-ideal MHD simulations. Additionally, the central protostar induces an effective net velocity field within the core which typically is not observed for prestellar cores. The core will continue to accumulate mass and as it becomes more and more dense, the random gas motions favor the direction of angular momentum.

3.2 Mass Accretion & Episodic behavior

The initial collapse of protostellar gas clouds begin with a given amount of angular momentum which in turn causes a rotation of the infalling gas and forms an accretion disk. A continuous infall of mass can sometimes lead to a disk that becomes unstable, resulting in a series of bursts

of mass during the accretion phase. Vorobyov & Basu [14] had modelled the evolution of a rotating magnetically supercritical cloud core up to the formation of a protostar and protostellar disk. The disk quickly experiences gravitational instabilities resulting in the formation of spiral arms with dense clumps arising from the infall of matter from the core's envelope. As the spiral arms evolve, the distribution of matter and angular momentum change. This disk fragmentation causes a series of bursts of mass accretion onto the central protostar resulting in enhanced accretion rates. Prior to this, the core experiences a rather steady change in the accretion rate.

4 Chapter Preview

Chapter 2 (published in the *Astrophysical Journal*, Bino and Basu [15]) begins with deploying the mathematical model first developed by Ewertowski & Basu (2013) [9] which results in a fully analytical expression for the magnetic field components in cylindrical coordinates. The single plane model is fit to the polarimetry data measured by Kandori et al [4] for the prestellar core FeSt 1-457. The model is independent of any assumed density profile for the core allowing for the deployment of more observationally-motivated density models. The robustness is back tested via a radiative transfer simulation used to produce synthetic polarization maps resulting from the integrated scattering and emission properties of the dust grains. Comparing these synthetic maps with the true polarization segments provides a more effective comparison over the single plane fit. Furthermore, the model enables direct estimate of the mass-to-flux ratio, which is used to study some of the mechanisms that govern the collapse.

Chapter 3 extends the evolutionary phase of prestellar cores to the early stages of protostellar formation. By utilizing non-ideal MHD simulations on a nested grid, one can resolve and model properties of protostellar cores on very small length scales. In doing such, it enables the modelling of the core's polarization states on these small length scales. By inputting the MHD data into the radiative transfer code POLARIS, a series of synthetic polarization maps can be simulated at all scales within the nested grid. This outputs high resolution polarization maps resolving the small scale regions that are typically difficult to measure. Furthermore, by matching length scales, the synthetic polarization maps are used to model the polarization state of Orion Source I. Through a series of declination angles and mean field analysis, the model can infer the emission or absorption properties of the observed dust grains.

Chapter 4 provides a study on the dynamics of the mass accretion properties in the late protostellar phase. With the introduction of machine learning practices, a neural network is demonstrated to providing an alternative model-free approach to predictive analysis. By simulating a series of hydrodynamic and stellar evolution models, an optimized form of the parallel echo state neural network (referred to as the OWA-ESN) is shown to extract critical underlying physical properties and temporal features in each simulation model. The methodology is introduced as a robust and efficient alternative to time series forecasting for data that is highly chaotic with no known governing model.

LARGE SCALE STRUCTURE – PRESTELLAR CORES

Key Words: stars: formation – ISM: magnetic fields – mathematical model

5 Introduction

Many properties governing prestellar cores are linked to the physical properties of the encompassing molecular cloud prior to dynamical collapse. The threading magnetic field is an important deterministic factor in the stability of the cloud [3, 4]. The cloud is said to be magnetically supercritical if the contraction mechanism is gravitationally dominated, dragging field lines inward with the direction of contraction. This scenario occurs rapidly on the dynamical time scale. Furthermore, the core is in a magnetically subcritical state if the contraction mechanism is magnetically driven due to the quasi-static evolution from the drift of neutral particles through ambipolar diffusion [16]. The subcritical scenario takes place on much longer time scales (so called ambipolar diffusion time scales [9]). The drift of neutral particles toward the center of the cloud slowly increases the mass to flux ratio M/Φ [4]. A cloud that is magnetically subcritical has a mass to flux ratio smaller than some defined critical ratio $(M/\Phi)_{\text{crit}}$ which is given for cases of perfectly balanced gravitational and magnetic contributions to contraction. Once $M/\Phi > (M/\Phi)_{\text{crit}}$ the cloud has now transitioned to a supercritical state which triggers the onset of dynamical collapse. A cloud that is initially in the supercritical state undergoes dynamical collapse sooner unless there is internal thermal and turbulent pressure for support [4]. This motivates further investigation of the magnetic properties in these prestellar cores. An estimate of the magnetic field morphology and potentially even magnetic field strength can be used to determine the magnetic state of the core in terms of its stability. Given that polarimetry can be used to infer the magnetic field morphology, it cannot be used to directly estimate magnetic field strength. This is due to there being many uncertainties in the properties of the dust grains and the respective efficiency in their alignment mechanisms. Direct measurement is possible through Zeeman effect [11], but has proven to be difficult due to measurement sensitivity and resolution. With the growing sample of polarization maps available, it is important to develop robust methods that can infer the magnetic properties and mass-to-flux ratios from data. A common method to model hourglass structures has been to fit a two dimensional single plane model to polarization segments. Girart et al. 2006 and Kandori et al. 2017 [4, 6] have done this utilizing a set of nested parabolas which although it provides a good fit visually, it is a purely mathematical approach and does not provide any direct physical information. Flux-freezing models [17–19] will infer a central magnetic field through column density estimates. The model of Ewertowski & Basu (2013) [9] assumes an axisymmetric system in cylindrical coordinates and is a self-consistent solution of Maxwell’s equations for a system with a current density that is confined near the equatorial plane in cylindrical geometry.

6 Mathematical Model

Modelling an hourglass magnetic field is the focal point of this chapter. A mathematical model [9] is derived and discussed for prestellar and protostellar cores. The model is built up from the fundamental equations of electromagnetism and is best suited for fitting polarization data.

6.1 Model Description

The model assumes a prestellar core is threaded with a background magnetic field in the z -direction. This background field will be distorted within the core. Let \mathbf{B}_0 denote the background field given as:

$$\mathbf{B}_0 = B_0 \hat{z}$$

and thus the total magnetic field can be written as:

$$(6.1) \quad \mathbf{B} = \mathbf{B}_c + \mathbf{B}_0$$

where \mathbf{B}_c is the local magnetic field generated within the core. The model assumes that the magnetic field is axisymmetric about the φ -direction and the magnetic field is taken as $B \equiv B(r, z)$. The assumption is based on the premise that the model accounts for the core on a large scale and that the twisting and irregularities in B_φ become critical only at the center of the core on a scale much smaller than that of the core itself. The model's magnetic field function is derived directly from Maxwell's equations and yields an explicit analytic expression for the magnetic field components along the r - and z -directions for a current density that is normally distributed along the vertical. This gives a magnetic vector potential function as:

$$(6.2) \quad A(r, z) = \sum_{m=1}^{\infty} k_m J_1(\sqrt{\lambda_m} r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} + \operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} \right]$$

for the coefficients:

$$(6.3) \quad k_m = \frac{2h\pi^{3/2} e^{h^2\lambda_m/4}}{cR^2\sqrt{\lambda_m} [J_2(\sqrt{\lambda_m}R)]^2} \int_0^R f(\xi) J_1(\sqrt{\lambda_m}\xi) \xi d\xi$$

The magnetic field can explicitly be found using the relation $\mathbf{B} = \nabla \times \mathbf{A}$ which yields the expressions:

$$(6.4) \quad B_r(r, z) = \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_1(\sqrt{\lambda_m} r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} - \operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} \right]$$

$$(6.5) \quad B_z(r, z) = \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_0(\sqrt{\lambda_m} r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} + \operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} \right] + B_0$$

where k_m , h , R and B_0 are taken as free parameters.

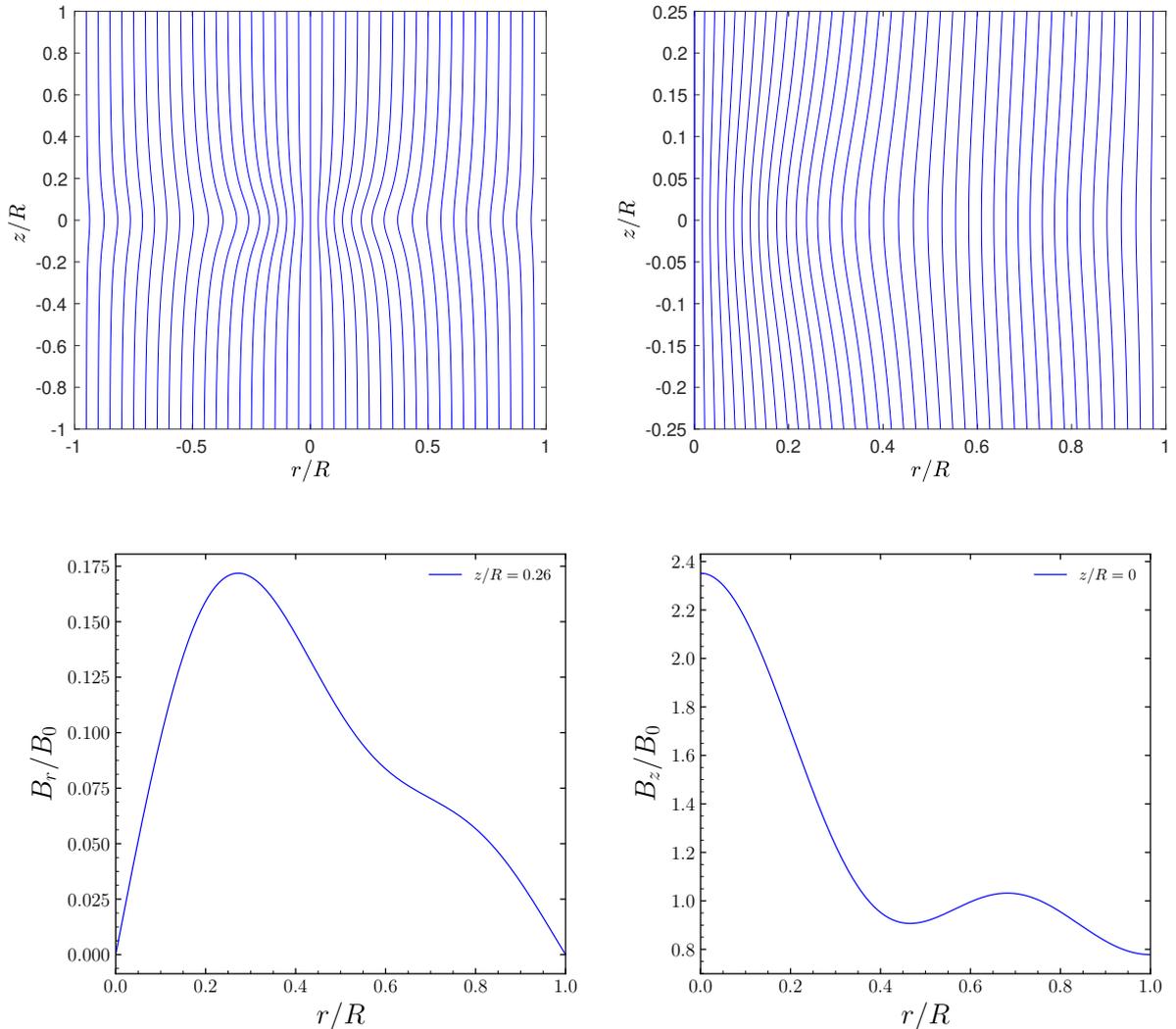


Figure 4: Magnetic field lines (**top**) and normalized field strength (**bottom**) using the parameters in the Basu and Ewertowski [9].

The reader is referred to appendix A1.1 for the complete mathematical derivation.

6.2 Parameters & Reformulation

It serves value to discuss the fitting parameters in the model, primarily the background field B_0 . In Basu and Ewertowski (2013) [9], they suggest that the background field should be independent to the fit. Measuring magnetic field directly is very difficult and thus more so than not, one turns to indirect methods to infer the magnetic field. Understanding how B_0 affects the hourglass pattern can give insight as to what the B_0 should be in practice. For demonstration purposes, the Basu, Ewertowski model parameters will be used and the background field will be varied. Looking at the following figures show that as the background field increases, the

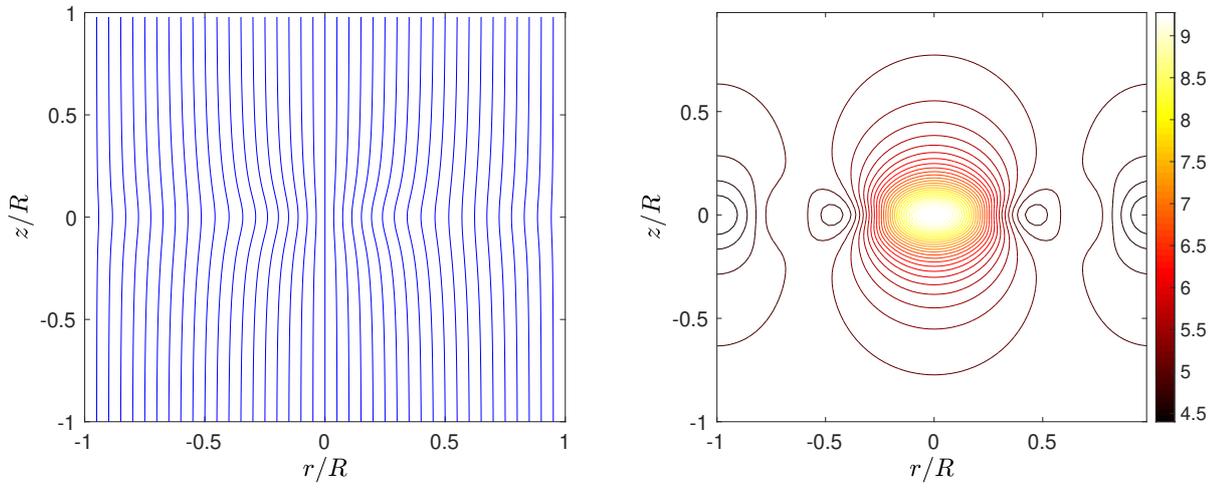


Figure 5: Background field $B = 1.5B_0$. **Left:** Magnetic field lines. **Right:** Contours of the total field strength.

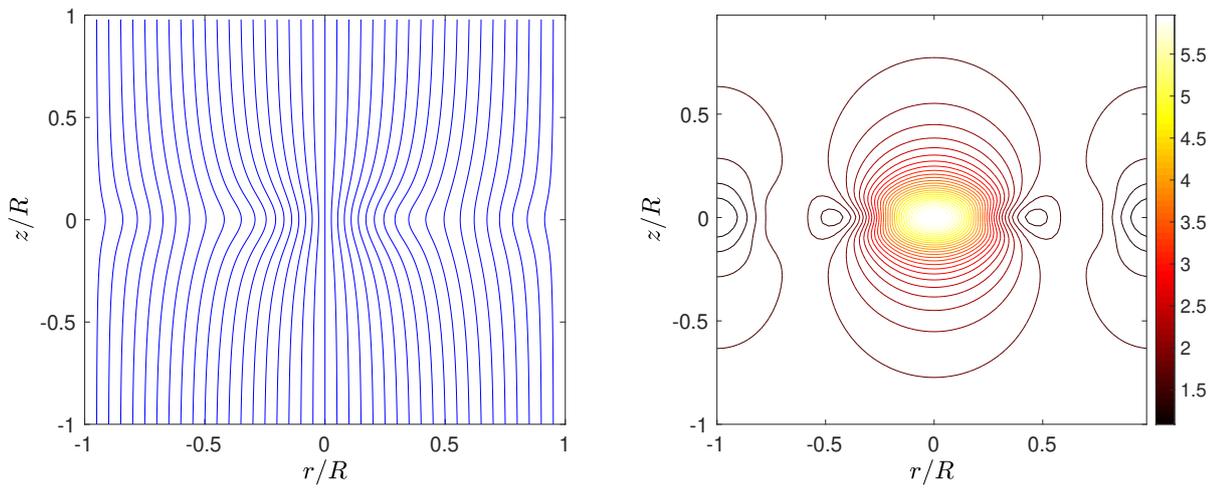


Figure 6: Background field $B = 0.5B_0$. **Left:** Magnetic field lines. **Right:** Contours of the total field strength.

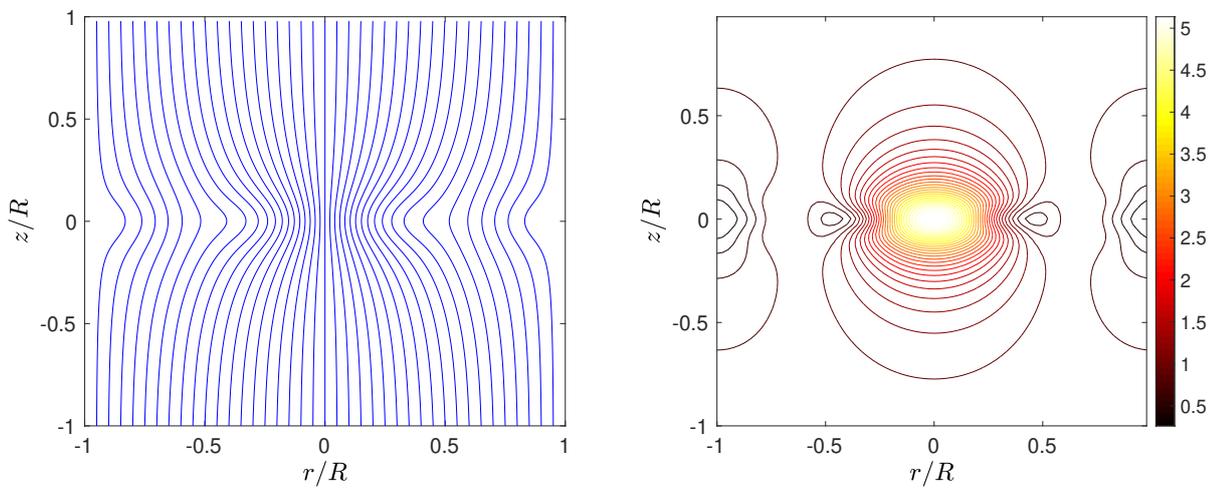


Figure 7: Background field $B = 0.25B_0$. **Left:** Magnetic field lines. **Right:** Contours of the total field strength.

hourglass pattern is masked and overtaken by B_0 . This is expected, but the more interesting property is what happens as the field strength is reduced. The bulk of the field strength is still concentrated at the core, but as $B_0 \rightarrow 0$ the hourglass 'pinch' at the core is amplified. The weaker the external field is, the more the field lines get distorted within the core. This can give an indication in practice as to how strong the relative background field is in regards to the field strength within the core. Naturally it would make more sense to make a fit to a model that is normalized to the background field. Since the background field is a difficult quantity to measure directly, one can reformulate 6.4 and 6.5 to being normalized to B_0 which would yield solutions for B_r and B_z normalized to the background field. This would not change anything in terms of fitting polarization data since polarization maps only give data for B_r/B_z and so a normalized model in principle still can be used to fit polarization data. Normalizing equations 6.4 and 6.5 involves scaling them to some benchmark quantity and expressing outputs in terms of said quantity. With that being said, define the following quantities normalized to their respective benchmark variables:

$$\begin{aligned}\tilde{r} &= r/R & \tilde{z} &= z/R & \eta &= h/R \\ \tilde{B}_r &= B_r/B_0 & \tilde{B}_z &= B_z/B_0\end{aligned}$$

A dimensional analysis is provided in appendix A1.2. The quantities are substituted into the model equations to yield:

$$\begin{aligned}\tilde{B}_r \times B_0 &= \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_1(\sqrt{\lambda_m} \tilde{r} R) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m} \eta R}{2} - \frac{\tilde{z} R}{\eta R} \right) e^{-\tilde{z} R \sqrt{\lambda_m}} - \operatorname{erfc} \left(\frac{\sqrt{\lambda_m} \eta R}{2} + \frac{\tilde{z} R}{\eta R} \right) e^{\tilde{z} R \sqrt{\lambda_m}} \right] \\ \tilde{B}_z \times B_0 &= \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_0(\sqrt{\lambda_m} \tilde{r} R) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m} \eta R}{2} + \frac{\tilde{z} R}{\eta R} \right) e^{\tilde{z} R \sqrt{\lambda_m}} + \operatorname{erfc} \left(\frac{\sqrt{\lambda_m} \eta R}{2} - \frac{\tilde{z} R}{\eta R} \right) e^{-\tilde{z} R \sqrt{\lambda_m}} \right] + B_0\end{aligned}$$

$$(6.6) \quad \tilde{B}_r = \sum_{m=1}^{\infty} \beta_m J_1(a_{m,1} \tilde{r}) \left[\operatorname{erfc} \left(\frac{a_{m,1} \eta}{2} - \frac{\tilde{z}}{\eta} \right) e^{-a_{m,1} \tilde{z}} - \operatorname{erfc} \left(\frac{a_{m,1} \eta}{2} + \frac{\tilde{z}}{\eta} \right) e^{a_{m,1} \tilde{z}} \right]$$

$$(6.7) \quad \tilde{B}_z = \sum_{m=1}^{\infty} \beta_m J_0(a_{m,1} \tilde{r}) \left[\operatorname{erfc} \left(\frac{a_{m,1} \eta}{2} + \frac{\tilde{z}}{\eta} \right) e^{a_{m,1} \tilde{z}} + \operatorname{erfc} \left(\frac{a_{m,1} \eta}{2} - \frac{\tilde{z}}{\eta} \right) e^{-a_{m,1} \tilde{z}} \right] + 1$$

where the dimensionless variable β_m has been defined to be:

$$\beta_m = \frac{k_m \sqrt{\lambda_m}}{B_0}$$

When presented with polarization data, one should now seek to fit the ratio $B_r/B_z \equiv \tilde{B}_r/\tilde{B}_z$ to the relative angles made with the vertical axis. The β_m terms suggest that B_0 and R can be selected at liberty, so long as the values for β_m remain unchanged. That is, in order to keep the model unchanged with respect to the fit, the quantities B_0 and R should scale inversely.

7 Polarimetry & Model Fit

Polarization data will be used to elucidate a magnetic field pattern along with other magnetic properties governing the core. The paramagnetic relaxation of dust grains has been proposed to give rise to magnetic alignment and under this proposal, polarimetry data becomes a trivial route to magnetic field modelling opposed to the direct measurements through Zeeman splitting. Equations 6.4 and 6.5 model the magnetic field directly in conjunction with the offset angles of each polarization vector.

7.1 Polarization Vectors

Polarization maps can serve as a valuable tool when studying a core's magnetic properties. The (at least partial) alignment of grains leads to dust emission with electric field perpendicular to the local magnetic field direction, or absorption of background starlight (dichroic extinction) that leaves a net polarization of the electric field that is parallel to the local magnetic field (see figure 8). Therefore, measurements of near infrared absorption will showcase polarization maps containing polarimetry vectors aligned parallel to the core's magnetic field.

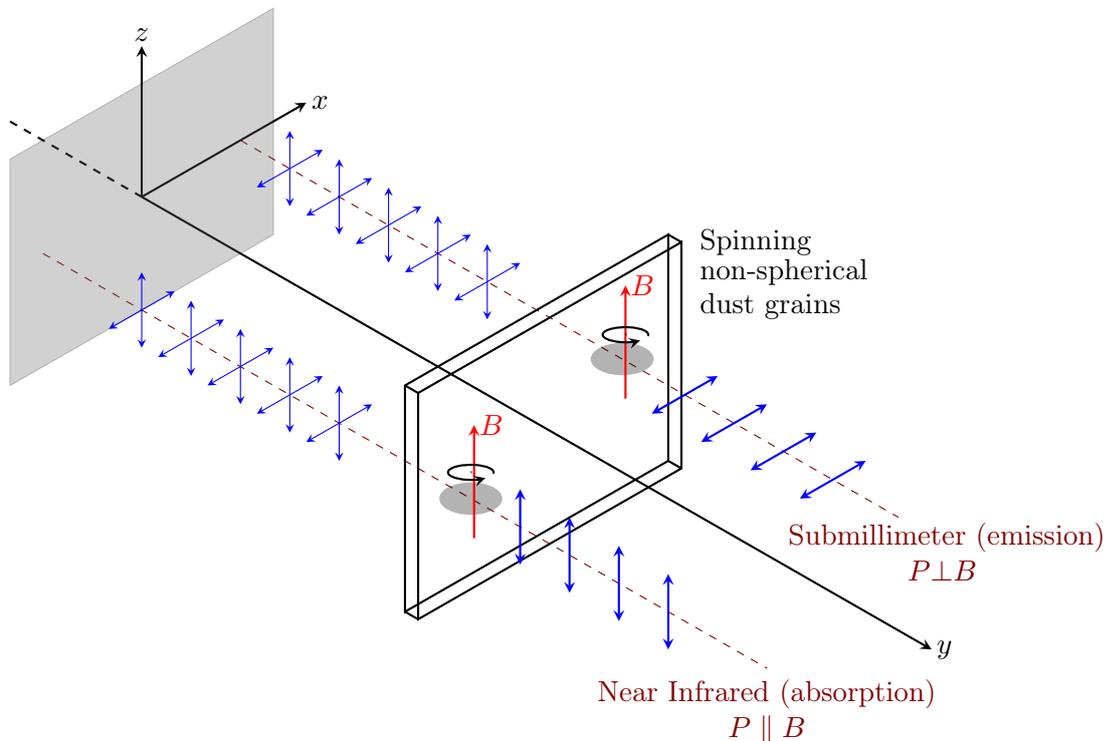


Figure 8: Unpolarized electromagnetic radiation emitted from a source in the $x - z$ plane will interact with spinning non-spherical dust grains and can be observed to have polarization vectors P oriented perpendicular to the local magnetic field for dust thermal emission or parallel to the local magnetic field for near infrared measurements.

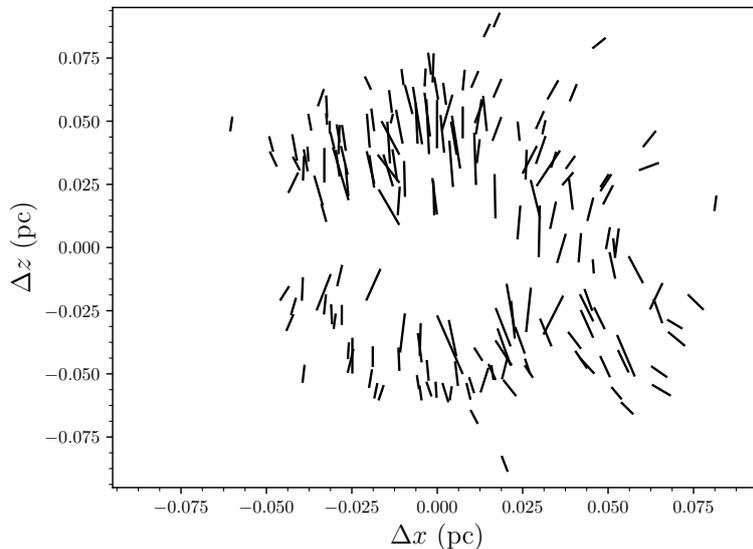


Figure 9: Polarization vectors for the core FeSt 1-457 [4]. Each dark line corresponds to polarization in the near-infrared due to dichroic extinction of light from a background star. The direction of each polarization vector is taken to be parallel to the magnetic field. The lengths represent the degree of polarization, with the longest line corresponding to about 5% polarization.

The simulation methods studied in section 8 model submillimeter emission due to thermal emission of dust grains and will result in polarization vectors oriented perpendicular to the magnetic field lines. Therefore, these alignment properties must be considered for simulation to observation comparisons. It should also be noted that while polarization measurements can reveal a magnetic field pattern, the polarization fraction cannot be used to directly measure the field strength due to the many uncertainties in dust grain properties and the alignment mechanism and efficiency. The field strength is traditionally estimated through the polarimetry using the Davis-Chandrasekhar-Fermi (DCF) method [8, 20] which estimates the magnetic field using Alfvénic fluctuations in conjunction with the density, velocity dispersion and the dispersion in the magnetic field lines. Rather than using the traditional DCF method, an estimate for magnetic field strength will be made directly using equations 6.4 and 6.5 in the model in combination with a background field estimate. Polarization maps are composed of vectors of polarized light where each vector makes an angle ϕ with the vertical axis and is related to the magnetic field through the tangent:

$$(7.1) \quad \tan \phi_i = \frac{B_r(r_i, z_i)}{B_z(r_i, z_i)}$$

Once the orientation angles have been sampled over the entire data set, the ratio B_r/B_z will be fit to 7.1 using some non-linear fitting method.

7.2 Polarimetry of FeSt 1-457

The observed data measures the near-infrared polarimetry of the dichroic extinction of the core through the radiation of roughly 185 background stars and is shown in figure 9. Kandori et al 2017 [4] modelled the magnetic field through a series of nested parabolas which although are capable of capturing the strong curvature of polarization vectors in the core's outer region, they provide little direct physical information regarding the core's magnetic field. The model that has been developed in section 6 overcomes this by defining an explicit expression for the components B_r and B_z . The data was extracted using *Plot Digitizer*, a software available to digitize scanned numerical data from figures. Matlab's SQP is used to fit the model magnetic field by minimizing the sum of squared residuals (the reader is referred to appendix A1.3 for details surrounding the minimization):

$$(7.2) \quad \min_{\beta, \eta} \|\Delta(r, z)\|^2 = \operatorname{argmin}_{\beta, \eta} \sum_i \left\| \frac{B_r(r_i, z_i)}{B_z(r_i, z_i)} - \tan \phi_i \right\|^2$$

Applying the algorithm and non-dimensionalizing the parameters gave:

$$\beta_1 = 15.3582, \quad \beta_2 = 0.5020, \quad \beta_3 = 23.4647, \quad \eta = 0.0366$$

The number of terms in the infinite series of the model magnetic field represent the number of modes used to describe the waveform of the magnetic field. Based off of an analysis testing higher order modes up to $m = 5$, it was found that these higher order modes ($m \geq 4$) contribute minimally to the fitting and in some cases reduce the model's performance (similar performance was realized in the original model implementation by Basu and Ewertowski [9]) and so the series has been truncated after 3 terms in the summation. An estimate for the core radius is taken

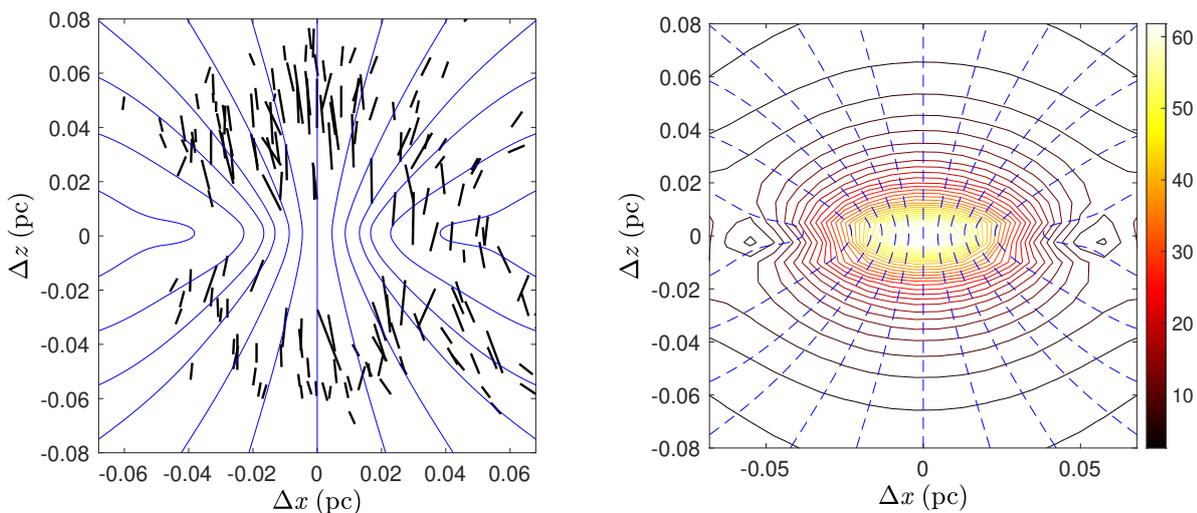


Figure 10: The model magnetic field lines taken at the midplane cut for $y = 0$ with optimal parameters fit to the polarimetry data of FeSt 1-457. **Left:** A streamline plot overlaid to the polarimetry data. **Right:** A streamline plot superimposed to the contours representing the total magnetic field magnitude of the model.

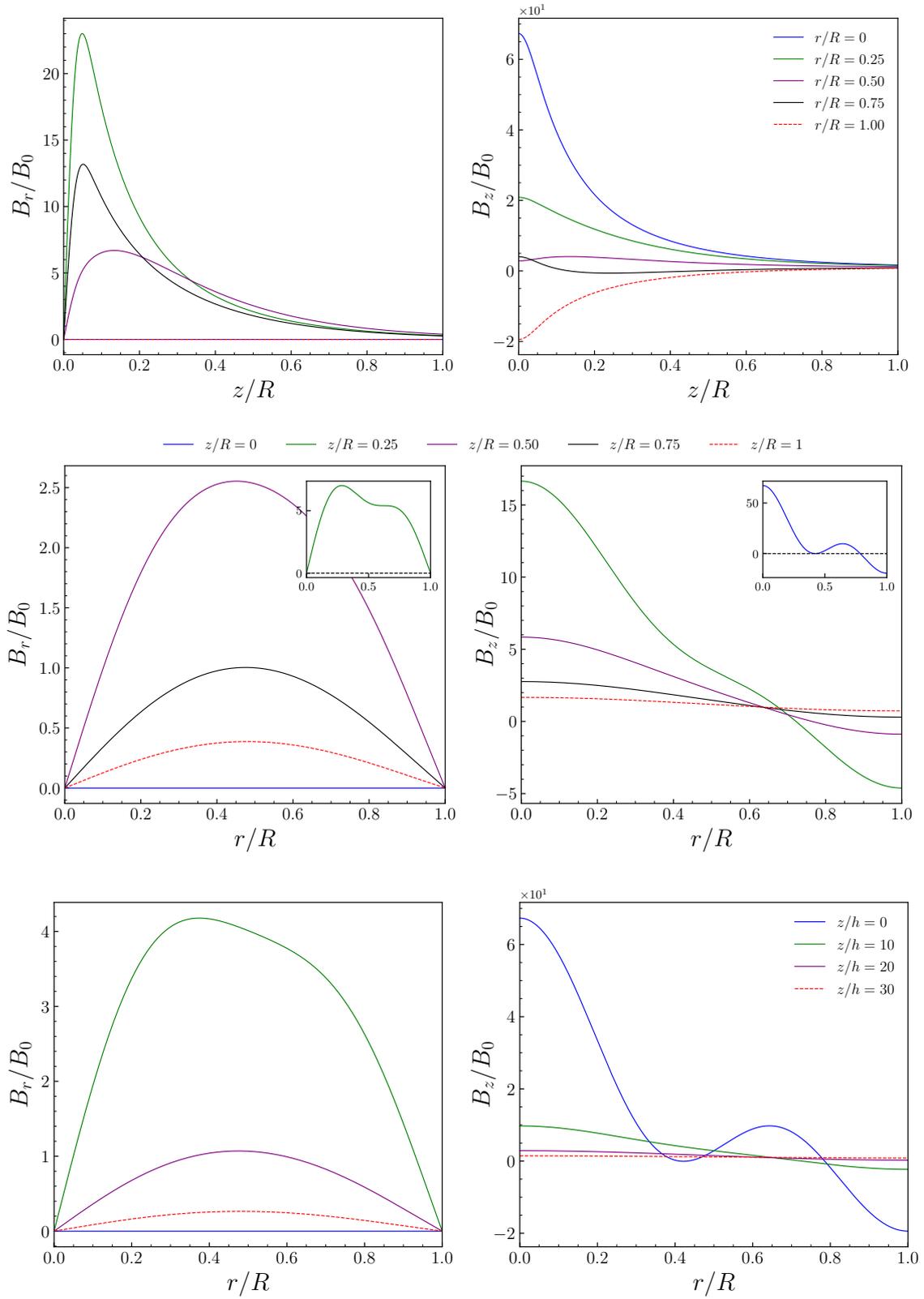


Figure 11: Model magnetic field plot for the full plane fit at various slices in the r & z -plane. **Left:** Plots of the r -component. **Right:** Plots of the z -component.

in agreement to a fit of the gas density by Kandori et al [21] with a radius $R = R_{\text{gas}} \approx 0.1$ pc. The left panel of figure 10 demonstrates contours to a total magnetic field strength that is estimated to be normalized to the background magnetic field B_0 . That is, the colorbar takes units of $|B|/B_0$. Figure 11 demonstrates the radial and vertical dependencies of \tilde{B}_r and \tilde{B}_z taken for various slices along r/R and z/h . The model's boundary conditions require $|B| \rightarrow B_0 \hat{z}$ as $z/R \rightarrow 1$ which implies that $B_r \rightarrow 0$ for $r \rightarrow \infty$ and $z \rightarrow \infty$. All panels of figure 11 demonstrate satisfying these conditions. Two new models are constructed from the conditional fit of the polarimetry to either plane of the data set. That is, because there are more polarization vectors in the right plane (to the right side of a vertical line along $z = 0$), two additional fits are performed individually for the right and left plane of the data. Applying the algorithm to the left plane of the data yields the optimal parameters:

$$\beta_1 = 0, \quad \beta_2 = 21.7963, \quad \beta_3 = 10.2505, \quad \eta = 0.0013$$

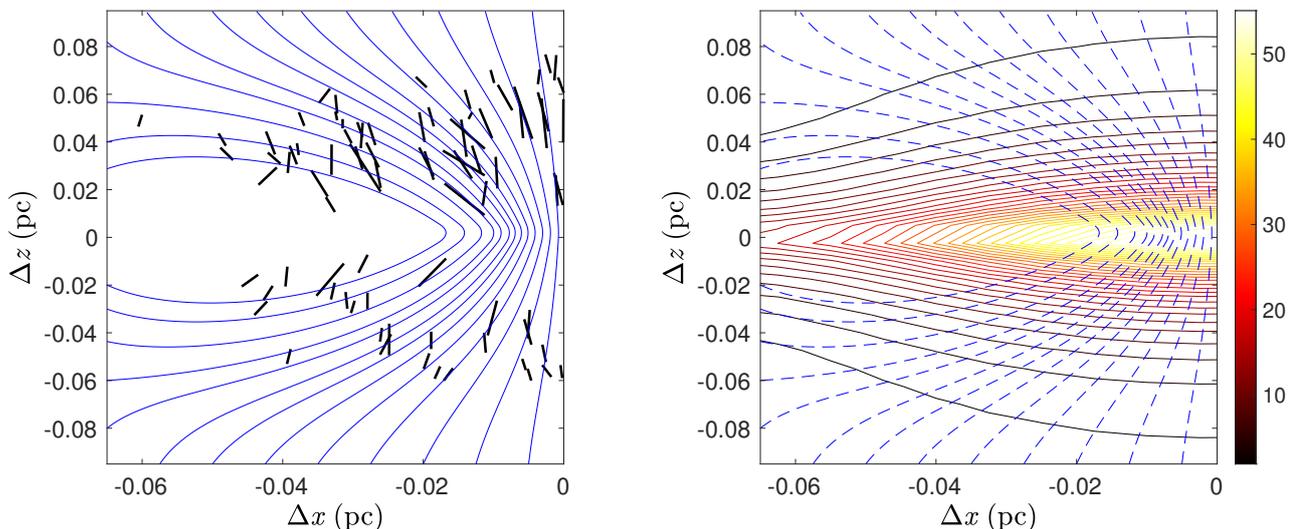


Figure 12: The model magnetic field lines taken at the midplane cut for $y = 0$ with optimal parameters fit to the left plane of the polarimetry data of FeSt 1-457. **Left:** A streamline plot overlaid to the polarimetry data. **Right:** A streamline plot superimposed to the contours representing the total magnetic field magnitude of the model.

The implied reconnecting loops in the field lines of figure 12 are indicative of a weak B_0 value. Figure 13 demonstrates different positional dependencies in the components \tilde{B}_r and \tilde{B}_z from figure 11, but none the less abides by the boundary conditions set forth in the model development. The relative magnetic field strength shown in the right panel of figure 12 is comparable to the original model and demonstrates similar distribution over the $x - z$ plane. Finally, a fit is performed to the model's right plane in similar fashion by applying the SQP algorithm yielding the solutions:

$$\beta_1 = 0 \quad \beta_2 = 4.5147 \times 10^3 \quad \beta_3 = 2.1968 \times 10^3 \quad \eta = 7.3099 \times 10^{-4}$$

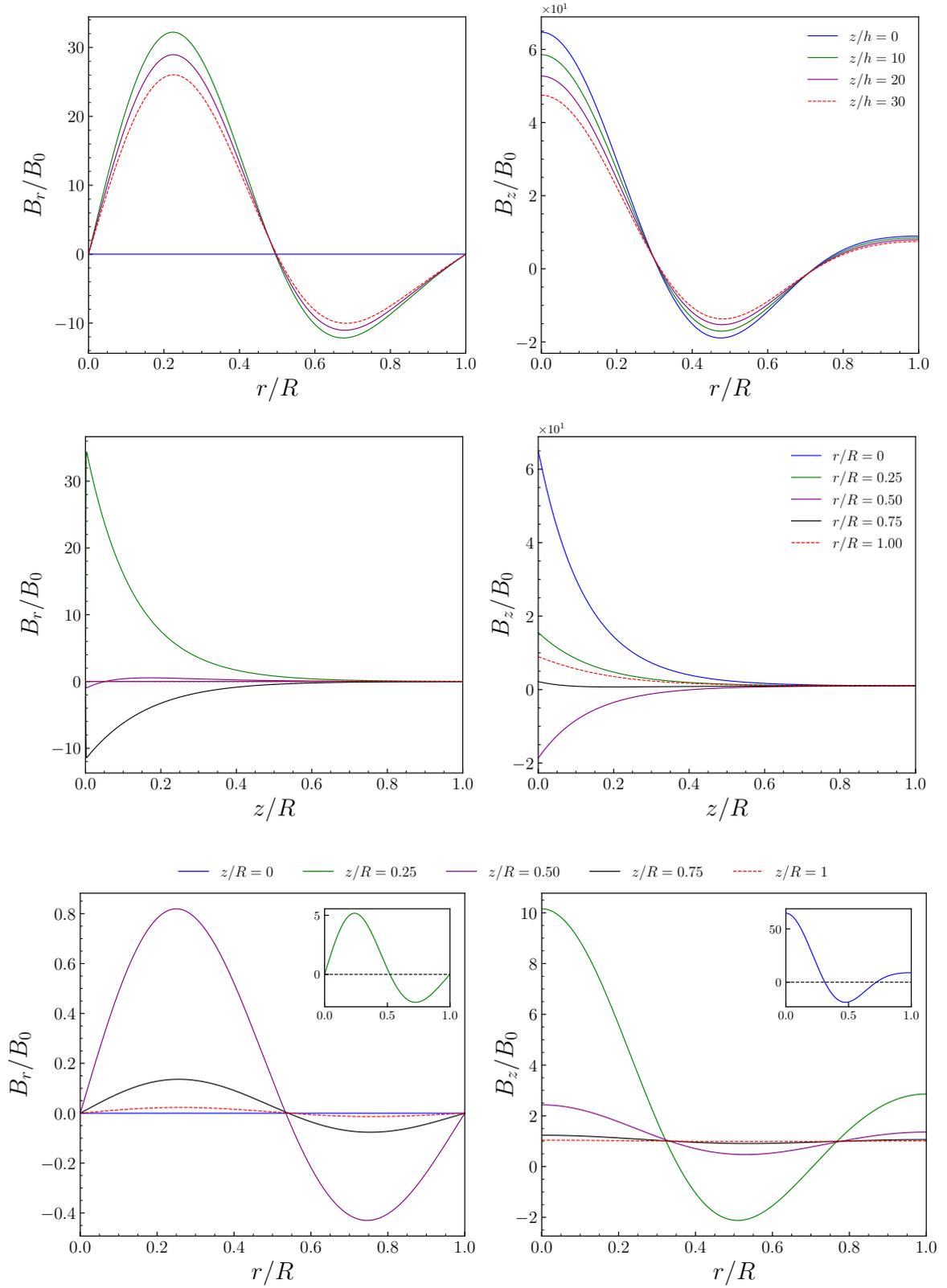


Figure 13: Model magnetic field plot for the left plane fit at various slices in the r & z -plane. **Left:** Plots of the r -component. **Right:** Plots of the z -component.

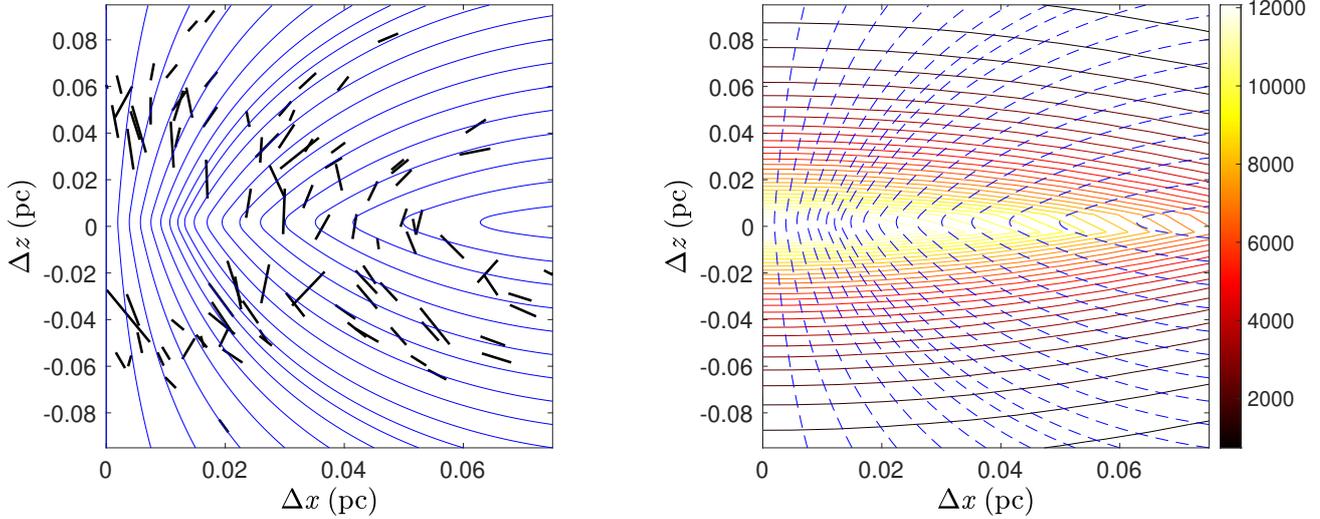


Figure 14: The model magnetic field lines taken at the midplane cut for $y = 0$ with optimal parameters fit to the right plane of the polarimetry data of FeSt 1-457. **Left:** A streamline plot overlaid to the polarimetry data. **Right:** A streamline plot superimposed to the contours representing the total magnetic field magnitude of the model.

The morphology in figure 14 shares similar properties in terms of how the magnetic field strength is distributed over the core, but favors a locally generated field that is orders of magnitude stronger than the background field. Figure 15 demonstrates similar component behavior of the magnetic field with respect to both positional arguments. Each model was fit to the magnetic field independently from the gas density profile, which advantageous to the flux freezing model in Kandori et al [22] which assumes a scaling of the form $|B| \propto \rho^\kappa$. Modelling the gas density separately is advantageous when running radiative transfer calculations as it allows for a more robust model to be studied through simulation. The magnetic field and density profiles need not be related to each other in a manner that assumes flux freezing, since some magnetic diffusion may have occurred during the core formation process. Additionally, a fit was performed by fixing the core radius to $R = 0.2$ pc, corresponding physically to a condensation of the core from a significantly larger volume than that implied from the density profile. The fit is performed in the same fashion implementing the SQP algorithm and the optimal dimensionless parameters are:

$$\beta_1 = 0.611 \quad \beta_2 = 0.000 \quad \beta_3 = 4.480 \quad \eta = 7.360 \times 10^{-4}$$

where the magnetic field lines are demonstrated in figure 16. This model is representative of the core at a potentially earlier contraction state and can be compared to the full plane fit demonstrated earlier having an inferred core radius of $R \approx 0.1$ pc. In doing so, the values of B_r and B_z are demonstrated in figure 17 at various cuts along the r and z -directions. Similarly comparing with figure 11, it can be seen that B_r approaches zero at both $r = R$ and again at

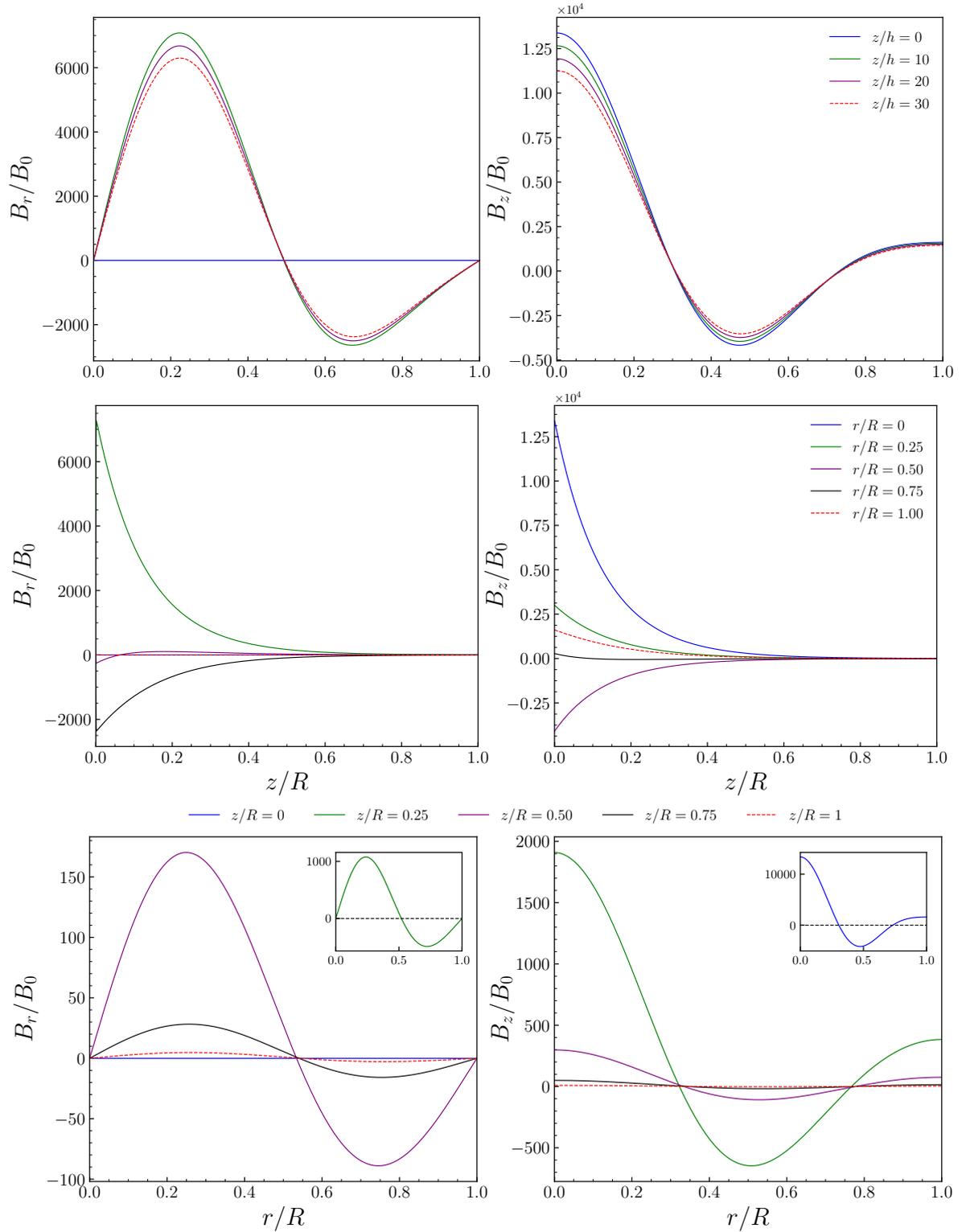


Figure 15: Model magnetic field plot for the right plane fit at various slices in the r & z -plane. **Left:** Plots of the r -component. **Right:** Plots of the z -component.

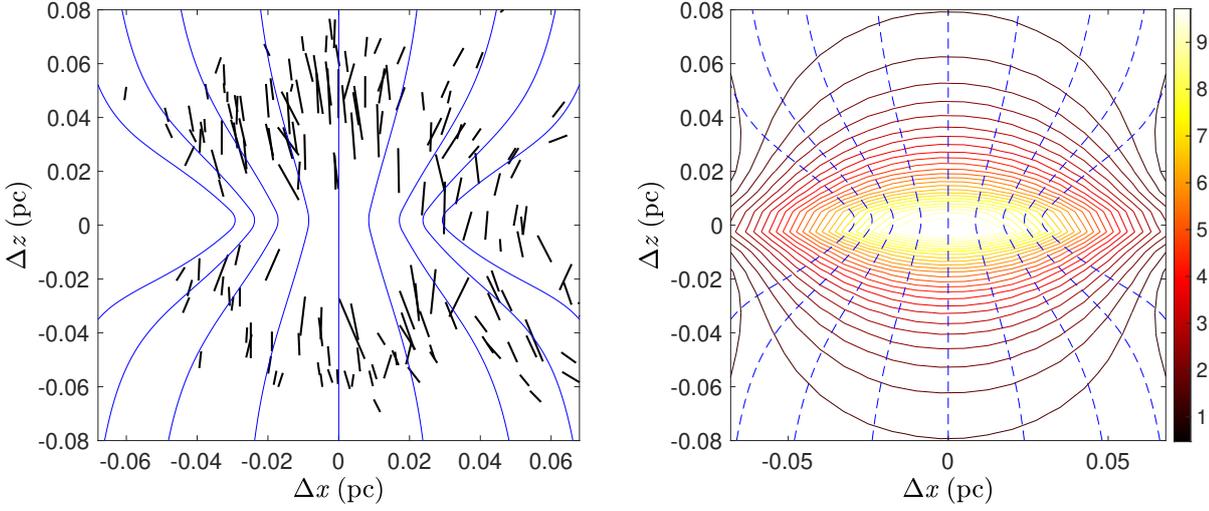


Figure 16: Model magnetic field fixing $R = 0.2$ pc taken at the midplane $y = 0$. The contours in the right plot represents the total magnetic field magnitude.

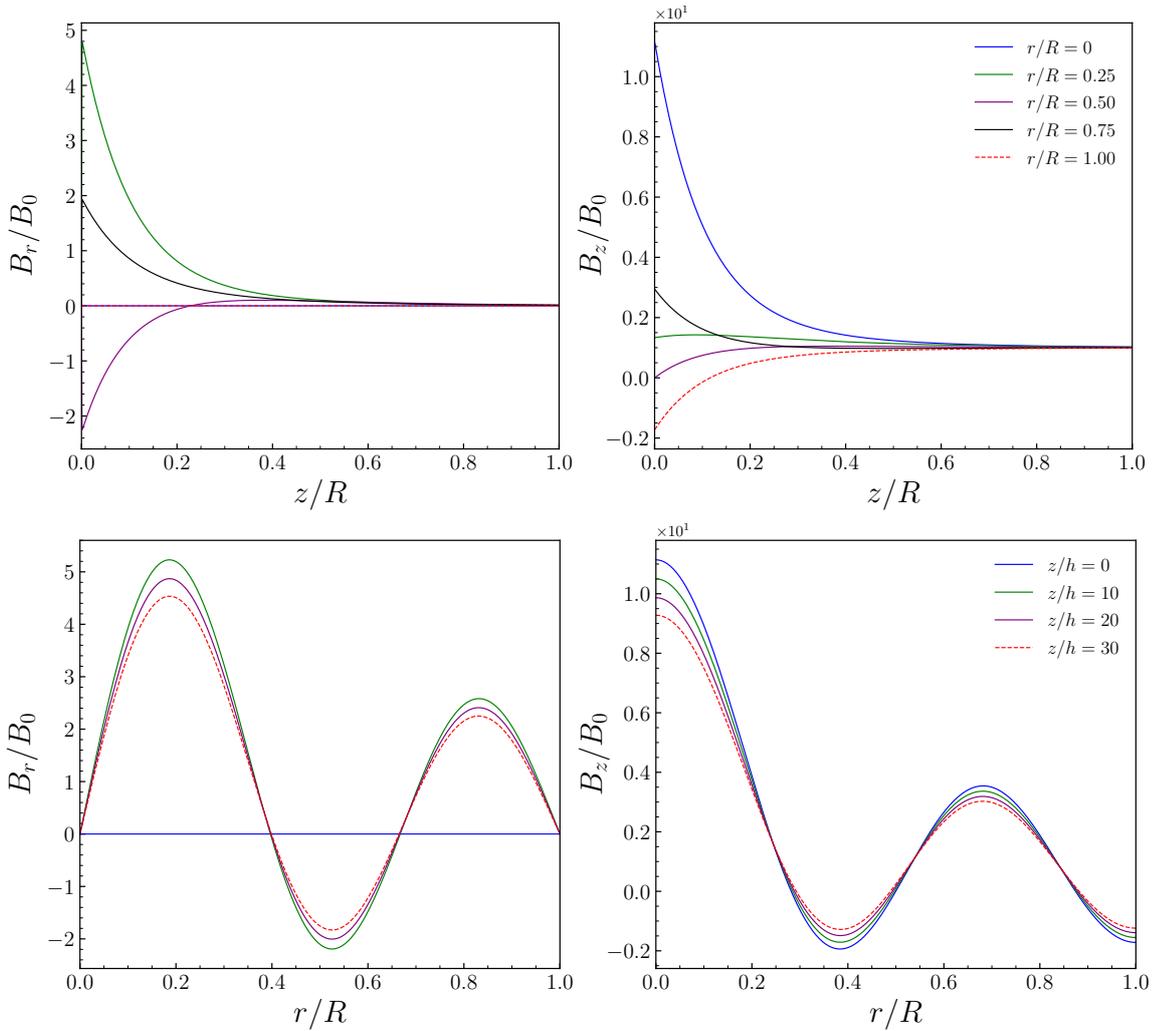


Figure 17: Model magnetic field plot for the model that fixes the core radius at $R = 0.2$ pc, at various slices in the r and z -planes. **Left:** Plots of the r -component. **Right:** Plots of the z -component.

large heights above the midplane. Strong field line curvature similarly is indicative of weak relative background magnetic field B_0 compared to the self-induced field. In fact, the dominant core generated field creates a region of closed field lines near the midplane about the region $0.05 \lesssim r \lesssim 0.1$ pc for both models. Physically, this means that when adopting $R = 0.2$ pc, the model allows the curved field lines to span a larger region than that associated with R_{gas} . Alternatively, the full plane fit which adopts $R = 0.1$ pc fits the entire field lines within the region enclosed by R_{gas} .

8 POLARIS Pipeline

This model makes use of the polarized radiation simulator POLARIS, a three-dimensional Monte Carlo radiative transfer code (see section 8.3). The code is written in C++ in an object oriented fashion but offers several extensions in Python via the *Polaris Tools* kit. POLARIS will be utilized to test the robustness of the magnetic field model for an environment where various astrophysical phenomena are included. These include the integrated scattering and emission properties of dust grains present within the core. The submillimeter dust thermal emission will have polarization vectors oriented along the direction of the electric field and perpendicular to the direction of local magnetic field. From this, the integrated magnetic field through the core can be inferred and cross validated with the morphology of the developed two dimensional model. The details discussed are for completeness, and a reader interested in the applications of POLARS should move ahead to section 9.

8.1 Stokes Vector and Polarization

The polarization state of light can be described by the Stokes parameters given as S_0 , S_1 , S_2 and S_3 . The parameters are better defined within the Stokes four-vector given as:

$$\mathbf{S} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$$

where the parameter I represents the total intensity, Q and U represent the state of linear polarization and V represents the circular polarization. The electric field as observed on the plane of the sky is represented as:

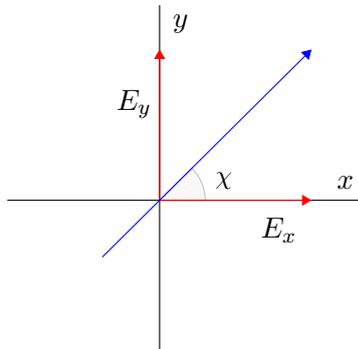


Figure 18: The electric field vector as seen from the plane of the sky where χ is the orientation angle.

This same electric field shares multiple representations which all serve purpose to redefining the Stokes vector in terms of electric field components given in three different coordinates in figure 19.

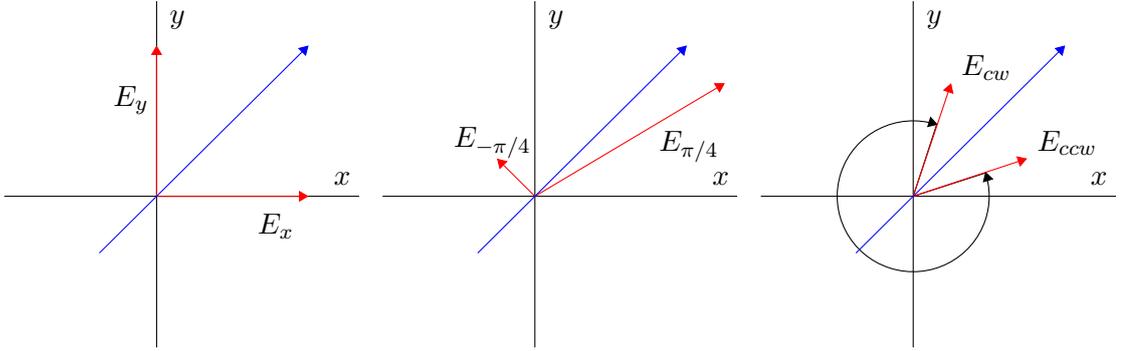


Figure 19: The same electric field vector represented in three different coordinates.

Polarized light is typically described by the orientation of oscillations in the coupled electric and magnetic fields. Polarization is typically classified as being either linear or elliptical. If the state of polarization is linear, the fields will oscillate in one direction, where elliptical polarization has the fields rotating in a plane as the wave propagates. The direction of rotation can be further classified as either right circular polarization or left circular polarization. At any fixed point in space the electric field vector will trace out an ellipse in the $x - y$ plane assuming the wave is travelling along z . A schematic image is demonstrated in figure 20. The spectral

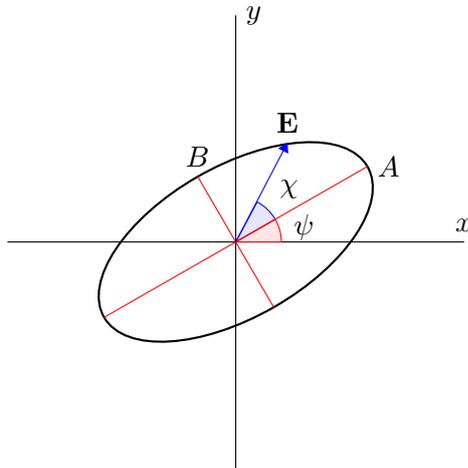


Figure 20: Schematic figure of the polarization ellipse. The semi-major and semi-minor axes of the ellipse have lengths A and B , respectively and the angle $\chi \equiv \omega t + \phi$.

decomposition formulism of the electric field is given as:

$$(8.1) \quad \mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi)$$

which can be decomposed into:

$$\begin{aligned} E_x(z, t) &= E_{x_0} \cos(\omega t - k_x z + \phi_x) \\ E_y(z, t) &= E_{y_0} \cos(\omega t - k_y z + \phi_y) \end{aligned}$$

Dividing by the amplitude and applying the double angle formula gives:

$$\begin{aligned}\frac{E_x}{E_{x_0}} &= \cos(\delta) \cos(\phi_x) - \sin(\delta) \sin(\phi_x) \\ \frac{E_y}{E_{y_0}} &= \cos(\delta) \cos(\phi_y) - \sin(\delta) \sin(\phi_y)\end{aligned}$$

where the parameter $\delta \equiv \omega t$ and without loss of generality assume $z = 0$, $E_x \equiv E_x(z, t)$ and $E_y \equiv E_y(z, t)$. The two expressions can be further rearranged and applying the double angle formula again gives:

$$\begin{aligned}\frac{E_x}{E_{x_0}} \sin(\phi_y) - \frac{E_y}{E_{y_0}} \sin(\phi_x) &= \cos(\delta) \sin(\phi_y - \phi_x) \\ \frac{E_x}{E_{x_0}} \cos(\phi_y) - \frac{E_y}{E_{y_0}} \cos(\phi_x) &= \sin(\delta) \sin(\phi_y - \phi_x)\end{aligned}$$

Squaring and adding the final two expressions results in the so called polarization ellipse equation:

$$(8.2) \quad \frac{E_x^2}{E_{x_0}^2} + \frac{E_y^2}{E_{y_0}^2} - \frac{2E_x E_y}{E_{x_0} E_{y_0}} \cos \hat{\phi} = \sin^2 \hat{\phi}$$

for $\hat{\phi} = \phi_y - \phi_x$. Taking the time average of equation 8.2 will represent the polarization in terms of measurables since the rotation angle ψ and the elliptic angle are not directly observable. The time average is given as:

$$(8.3) \quad \frac{\langle E_x^2 \rangle}{E_{x_0}^2} + \frac{\langle E_y^2 \rangle}{E_{y_0}^2} - \frac{2\langle E_x E_y \rangle}{E_{x_0} E_{y_0}} \cos \hat{\phi} = \sin^2 \hat{\phi}$$

where $\langle g(t) \rangle$ is:

$$\langle g(t) \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau g(t) dt$$

Multiplying equation 8.3 by $4E_{x_0}^2 E_{y_0}^2$ and computing the average gives:

$$2E_{x_0}^2 E_{y_0}^2 + 2E_{x_0}^2 E_{y_0}^2 - (2E_{x_0} E_{y_0} \cos \hat{\phi})^2 = (2E_{x_0} E_{y_0} \sin \hat{\phi})^2$$

Adding and subtracting $E_{x_0}^4 + E_{y_0}^4$ yields perfect squares and so:

$$(8.4) \quad (E_{x_0}^2 + E_{y_0}^2)^2 - (E_{x_0}^2 - E_{y_0}^2)^2 - (2E_{x_0} E_{y_0} \cos \hat{\phi})^2 = (2E_{x_0} E_{y_0} \sin \hat{\phi})^2$$

Using $S_0^2 = S_1^2 + S_2^2 + S_3^2$ and comparing with equation 8.4 gives the relation:

$$\begin{aligned}S_0 &= E_{x_0}^2 + E_{y_0}^2 \\ S_1 &= E_{x_0}^2 - E_{y_0}^2 \\ S_2 &= 2E_{x_0} E_{y_0} \cos \hat{\phi} \\ S_3 &= 2E_{x_0} E_{y_0} \sin \hat{\phi}\end{aligned}$$

The Stokes vector is now defined as:

$$\mathbf{S} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} E_{x_0}^2 + E_{y_0}^2 \\ E_{x_0}^2 - E_{y_0}^2 \\ 2E_{x_0}E_{y_0} \cos \hat{\phi} \\ 2E_{x_0}E_{y_0} \sin \hat{\phi} \end{bmatrix}$$

If the wave is represented in terms of complex amplitudes for planar waves:

$$(8.5) \quad \mathbf{E}(\mathbf{r}) = \mathbf{E}_0 e^{i\omega t}$$

then the Stokes parameters will be given as:

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} E_x E_x^* + E_y E_y^* \\ E_x E_x^* - E_y E_y^* \\ E_x E_y^* + E_y E_x^* \\ i(E_x E_y^* - E_y E_x^*) \end{bmatrix}$$

This representation can be further written in terms of the electric field components in figure 19 by defining:

$$E_{\pi/4} = \frac{E_x + E_y}{\sqrt{2}}, \quad E_{-\pi/4} = \frac{E_x - E_y}{\sqrt{2}}, \quad E_{cw} = \frac{E_x - iE_y}{\sqrt{2}}, \quad E_{ccw} = \frac{E_x + iE_y}{\sqrt{2}}$$

which results in

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} |E_x|^2 + |E_y|^2 \\ |E_x|^2 - |E_y|^2 \\ |E_{\pi/4}|^2 + |E_{-\pi/4}|^2 \\ |E_{cw}|^2 + |E_{ccw}|^2 \end{bmatrix}$$

The Poincaré sphere relates parameters to polarization measurements as depicted in figure 21.

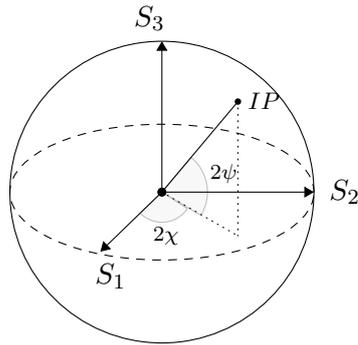


Figure 21: Poincaré sphere spanning S_1 , S_2 and S_3 representing the state of polarized light.

The sphere spans S_1 , S_2 and S_3 and using spherical coordinate relations, the Stokes parameters are given as:

$$S_0 = IP, \quad S_1 = IP \cos 2\chi \cos 2\psi, \quad S_2 = IP \sin 2\chi \cos 2\psi, \quad S_3 = IP \sin 2\psi$$

which define the relations:

$$P = \frac{\sqrt{Q^2 + U^2 + V^2}}{I}, \quad \frac{U}{Q} = \tan 2\chi$$

where P is the total degree of polarization based on light intensity. Furthermore, the linear and circular polarization are respectively taken as:

$$P_l = \frac{\sqrt{Q^2 + U^2}}{I} \in [0, 1], \quad P_c = \frac{V}{I} \in [-1, 1]$$

8.2 Radiative Transfer Equation

The general radiative transfer equation is given as:

$$\frac{d}{dl} \mathbf{S} = -\hat{\mathbf{R}}(\alpha) \hat{\mathbf{K}} \hat{\mathbf{R}}(\alpha)^{-1} \mathbf{S} + \mathbf{J}$$

where \mathbf{J} is the energy transfer contribution due to emission, $\hat{\mathbf{K}}$ is the Muller matrix describing the extinction, absorption, and scattering of the radiation interacting with the gas species and $\hat{\mathbf{R}}(\alpha)$ is the rotation matrix given as:

$$\hat{\mathbf{R}}(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(2\phi) & -\sin(2\phi) & 0 \\ 0 & \sin(2\phi) & \cos(2\phi) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The radiative transfer equation is solved for all Stokes parameters I , Q , U , V and in its most general form is given as:

$$\frac{d}{dl} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = - \begin{bmatrix} \alpha_I & \alpha_Q & \alpha_U & \alpha_V \\ \alpha_Q & \alpha_I & \kappa_V & \kappa_U \\ \alpha_U & -\kappa_V & \alpha_I & \kappa_Q \\ \alpha_V & -\kappa_U & -\kappa_Q & \alpha_I \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} + \begin{bmatrix} J_I \\ J_Q \\ J_U \\ J_V \end{bmatrix}$$

The radiative transfer equation is solved numerically by discretizing the derivative terms over a discrete mesh. POLARIS uses a Runge-Kutta Fehlberg method as an integrator for the ODE which effectively computes fourth and fifth order RK schemes and compares the solutions to a given tolerance. The solution is accepted if the tolerance is met and rejected otherwise, resulting in a reduction of the step size h . In an effort to maintain the program's computational efficiency, the step may also be increased if solutions agree to a given number of significant digits. Each step requires the following indices:

$$k_1 = hf(t_k, y_k)$$

$$k_2 = hf\left(t_k + \frac{1}{4}h, y_k + \frac{1}{4}k_1\right)$$

$$\begin{aligned}
k_3 &= hf \left(t_k + \frac{3}{8}h, y_k + \frac{3}{32}k_1 + \frac{9}{32}k_2 \right) \\
k_4 &= hf \left(t_k + \frac{12}{13}h, y_k + \frac{1932}{2197}k_1 - \frac{7200}{2197}k_2 + \frac{7296}{2197}k_3 \right) \\
k_5 &= hf \left(t_k + h, y_k + \frac{439}{216}k_1 - 8k_2 + \frac{3680}{513}k_3 - \frac{845}{4104}k_4 \right) \\
k_6 &= hf \left(t_k + \frac{1}{2}h, y_k - \frac{8}{27}k_1 + 2k_2 - \frac{3544}{2565}k_3 + \frac{1859}{4104}k_4 - \frac{11}{40}k_5 \right)
\end{aligned}$$

for the fourth y_{k+1} and fifth order \hat{y}_{k+1} schemes given as:

$$\begin{aligned}
y_{k+1} &= y_k + \frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4101}k_4 - \frac{1}{5}k_5 \\
\hat{y}_{k+1} &= y_k + \frac{16}{135}k_1 + \frac{6656}{12825}k_3 + \frac{28561}{56430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6
\end{aligned}$$

The error is finally calculated:

$$\epsilon = \frac{|\hat{y}_{k+1} - y_{k+1}|}{\epsilon_{\text{tol}}|\hat{y}_{k+1} - \epsilon_{\text{abs}}|}$$

where ϵ_{tol} is the threshold tolerance of comparing the fourth and fifth order schemes and ϵ_{abs} is a threshold below which the value of the given solution component becomes unimportant. These values are taken to be $\epsilon_{\text{tol}} = 10^{-8}$ and $\epsilon_{\text{abs}} = 10^{-30}$. The step size h is adjusted according to:

$$h_{i+1} = \min(0.9h_i\epsilon^{-0.2}, 0.25h_i)$$

8.3 Monte Carlo Photon Transfer

Simulating photon transport using Monte Carlo methods involves expressing photon trajectories in terms of probability distributions which describe the step size of photon movement within some three dimensional grid made of cells. POLARIS runs the simulation using photon packets rather than individual photons to upkeep computational efficiency. Each packet comes from a source given some initial direction, energy per unit time, polarization state and one particular wavelength. Within each cell there is a radiation-dust interaction probability which depends on the path length l through the cell, the extinction cross section C_{ext} (total losses of energy due to both absorption and scattering), and the number density of dust within that cell n_d . The optical depth is defined through the statistical function:

$$\tau_s = -\ln(1 - X)$$

where $X \in [0, 1)$ is a sampled random number that is uniformly distributed. POLARIS defines one path length being the distance between two walls within one cell. The total optical depth a photon packet accumulates from all path lengths over successive cells is given as:

$$\tau_{\text{ext}} = \sum_{i=1}^N C_{\text{ext}}^{(i)} n_d^{(i)} l_i$$

where the extinction cross section, number density of dust and path length are all taken locally at the i -th cell. If

$$\tau_{\text{ext}} < \tau_s$$

the photon packet will interact with dust within the cell and the position will be adjusted to ensure $\tau_{\text{ext}} = \tau_s$. This interaction can be either scattering or absorption and emission. The likelihood of a scattering interaction occurring is given by a dust grain albedo defined at a given wavelength:

$$(8.6) \quad p_a(\lambda) = \frac{C_{\text{sca},\lambda}}{C_{\text{sca},\lambda} + C_{\text{abs},\lambda}}$$

taken to be $\in [0, 1]$. The terms $C_{\text{sca},\lambda}$ and $C_{\text{abs},\lambda}$ are the dust cross sections of scattering and absorption respectively. These cross sections are proportional to the physical cross section of the dust grain itself (related to the grain radius a), with some interaction efficiency ε :

$$\begin{aligned} C_{\text{sca}} &= \varepsilon_{\text{sca}} \pi a^2 \\ C_{\text{abs}} &= \varepsilon_{\text{abs}} \pi a^2 \end{aligned}$$

If $X < p_a$ the packet is absorbed and immediately re-emitted to sustain local thermodynamic equilibrium where the direction of re-emission followed is isotropic with a new wavelength sampled from:

$$p(\lambda, T_d) = \frac{\int_{\lambda}^{\lambda+d\lambda} C_{\text{abs},\lambda} D_T[B_{\lambda}(T_{d,i})] d\lambda}{\int_{\lambda_{\text{min}}}^{\lambda_{\text{max}}} C_{\text{abs},\lambda} D_T[B_{\lambda}(T_{d,i})] d\lambda}$$

where $D_T \equiv d/dT$, T_d is the dust temperature and $B_{\lambda}(T_{d,i})$ is the Planck function. If the dust temperature is unchanged then the wavelength is sampled from:

$$p(\lambda, T_d) = \frac{\int_{\lambda}^{\lambda+d\lambda} C_{\text{abs},\lambda} D_T[B_{\lambda}(T_d)] d\lambda}{\int_0^{\infty} C_{\text{abs},\lambda} D_T[B_{\lambda}(T_d)] d\lambda}$$

If $X > p_a$ the packet will scatter on the dust grain at some angle ψ . The scattering angle depends on the model's characteristic phase function $\Psi(\psi)$. The phase function describes the angular distribution of scattering on a grain at a given wavelength. Isotropic scattering assumes that the packet has equal probability of scattering in any given direction. A new direction \mathbf{r}' is given from sampling points that are uniformly distributed on a unit sphere having three parameters:

$$\begin{aligned} u &= 1 - 2X_1 \\ t &= 2\pi X_2 \\ r &= \sqrt{1 - u^2} \end{aligned}$$

where X_1 and X_2 are random numbers and the new direction is taken to be:

$$\mathbf{r}' = \begin{bmatrix} r \cos(t) \\ r \sin(t) \\ u \end{bmatrix}$$

If the dust grains are anisotropic having a preferential scattering angle, then POLARIS samples ψ from the Henyey-Greenstein (HG) phase function:

$$(8.7) \quad \Psi(\psi) = \frac{1}{4\pi} \frac{1 - g^2}{[1 + g^2 - 2 \cos(\psi)]^{3/2}}$$

The HG phase function benefits from a single parameter variation $g \in [-1, 1]$ allowing for backward, forward and isotropic scattering where:

$$\begin{aligned} g = 0, & \quad \text{Isotropic Scattering} \\ g = 1, & \quad \text{Forward Scattering} \\ g = -1, & \quad \text{Backward Scattering} \end{aligned}$$

The parameter g is pre-calculated by POLARIS based off a list of dust properties such as the grain material, size and wavelength.

8.4 Grain Alignment

To directly model the magnetic field in this chapter, perfect alignment is enforced in order to trace out the field lines from dust grain orientation. This will have some consequences that should be discussed when simulating polarization maps. The main observation being that the total amount of radiation being polarized will be much higher than what is typically observed in practice. That is because grains align perfectly to the perpendicular direction of the magnetic field, the efficiency of polarized light is much higher. In practice, grains are typically not perfectly aligned and this imperfect alignment property is used in chapter 3.

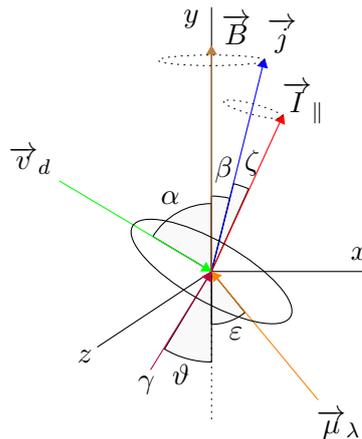


Figure 22: Grain's disrupted alignment with the magnetic field.

In the case of misaligned dust grains, the angular momentum vector \mathbf{j} will precess about the magnetic field direction causing a misalignment between the normal axis of the grain and the direction of the magnetic field. The above figure shows the configuration of a dust grain that

is imperfectly aligned with the magnetic field. The angular momentum precesses around the magnetic field and sweeps a cone with angle β about the magnetic field. It's important to note that the maximum moment of inertia vector I_{\parallel} also precesses about \mathbf{j} which is represented by the angle ζ . The electromagnetic wave is incident to the grain at an angle ϑ with the magnetic field. The figure also demonstrates disruption due to a net drift velocity \mathbf{v}_d and the radiation field $\boldsymbol{\mu}_\lambda$ which make angles α and ε with the magnetic field respectively. This disruption causes a reduction in polarization and can be quantified by calculating the Rayleigh reduction factor:

$$R = \langle G(\cos^2 \beta)G(\cos^2 \zeta) \rangle \in [0, 1]$$

where $G(x) = 1.5x - 0.5$. Grain alignment and polarization are determined by the first order moment $\langle \cos^2 x \rangle$ weighted over some distribution function. This is given as:

$$\langle \cos^2 x \rangle = \int f(x) \sin(x) dx$$

where $f(x)$ is taken as the distribution function and is dependant on the grain alignment mechanism selected. If $R = 1$ the polarization is at a maximum and if $R = 0$ the polarization is completely suppressed. POLARIS assumes that the precession angles are independent and so

$$\langle G(X)G(Y) \rangle \sim \langle G(X) \rangle \times \langle G(Y) \rangle \times (1 + f_c)$$

where f_c is the correlation factor (default $f_c = 0.6$) which dictates the correlation between $\langle G(X) \rangle$ and $\langle G(Y) \rangle$. This chapter enforces perfect alignment between magnetic field and dust grains and so will not utilize an alignment theories. Chapter 3 however, will use the Davis-Greenstein (DG) alignment mechanism offered by POLARIS. DG alignment theory [8] suggests that the disturbed alignment of \mathbf{j} with magnetic field \mathbf{B} comes about from external random collisions with surrounding gas. Paramagnetic relaxation in the material causes a torque that rotates the grain about the shorter axis. The alignment is determined by the parameter δ_0

$$\delta_0 = 2.07 \times 10^{20} \frac{\mathbf{B}^2}{n_g T_d \sqrt{T_g}}$$

which is the upper threshold for grain alignment. Grains with effective radius larger than δ_0 are considered to not significantly contribute to the net polarization. The distribution of the opening angle β is given analytically from Spitzer Jr and McGlynn [23]:

$$\langle \cos^2 \beta \rangle = \left(1 - \sqrt{\frac{\xi^2}{1 - \xi^2}} \right) \arcsin \frac{\sqrt{1 - \xi^2}}{1 - \xi^2}$$

where:

$$\xi^2 = \frac{a + \delta_0 \times T_d/T_g}{a + \delta_0}$$

which is a function of the grain size a and is effectively used to calculate the reduction factor R . POLARIS also offers other grain alignment theories briefly summarized as

-
- **Internal imperfect alignment:** Due to internal thermal fluctuations of the dust grains giving rise to precessions in the angular momentum vector about the magnetic field [24].
 - **Gold magneto mechanical alignment:** Due to a mechanical process through the interaction between the grains and a gaseous flux [24, 25].
 - **Radiative torque alignment:** Considers gaseous flux interactions to be random by which the effects are balanced by both the radiative pressure and the paramagnetic properties of the grains. Grains with $a > \lambda/2$ experience a differential torque on all facets due to left and right circular polarization [26].

8.5 Intensity and Polarization Maps

Non-spherical spinning dust grains will have the major axis align perpendicular to the local magnetic field. The grains will thus block more light along this direction and the re-emission will result in light that is polarized perpendicular to the local magnetic field. The Stokes vector can be rotated from the lab frame into the reference frame of the alignment direction to remove the extinction coefficients [27] $\alpha_U = \alpha_V = \kappa_U = \kappa_V = 0$ and the emissivity coefficients $J_U = J_V = 0$. The radiative transfer equation is therefore fully described by:

$$(8.8) \quad \frac{d}{dl} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = - \begin{bmatrix} \alpha_I & \alpha_Q & 0 & 0 \\ \alpha_Q & \alpha_I & 0 & 0 \\ 0 & 0 & \alpha_I & \kappa_Q \\ 0 & 0 & -\kappa_Q & \alpha_I \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} + \begin{bmatrix} J_I \\ J_Q \\ 0 \\ 0 \end{bmatrix}$$

For non-spherical grains, the cross section coefficients are described in terms of the grain's major axis x and minor axis y . The components of the extinction cross sections are given by:

$$C_{\text{ext},x} = \langle C_{\text{ext}} \rangle + \frac{1}{3}R \left(C_{\text{ext},||} - C_{\text{ext},\perp} \right)$$

$$C_{\text{ext},y} = \langle C_{\text{ext}} \rangle + \frac{1}{3}R \left(C_{\text{ext},||} - C_{\text{ext},\perp} \right) \left(1 - 3 \sin^2 \vartheta \right)$$

where $C_{\text{ext},||}$ and $C_{\text{ext},\perp}$ are the components parallel and perpendicular to the magnetic field. Here, $\langle \cdot \rangle$ is representative of the average taken over the orientation and for $\langle C_{\text{ext}} \rangle$ this is taken as:

$$\langle C_{\text{ext}} \rangle = \frac{2C_{\text{ext},||} + C_{\text{ext},\perp}}{3}$$

The cross section of extinction and linear polarization are given respectively as:

$$C_{\text{ext}} = \frac{C_{\text{ext},x} + C_{\text{ext},y}}{2}$$

$$C_{\text{ext,pol}} = \frac{C_{\text{ext},x} - C_{\text{ext},y}}{2}$$

The cross sections for scattering C_{sca} , absorption C_{abs} and $C_{\text{abs,pol}}$ and circular polarization $C_{\text{circ,pol}}$ are calculated in the same fashion. Therefore, components to the matrix equation

8.8 are taken as the cross sections weighted over the grain size distribution $n_d(a)$ between a minimum and maximum grain size a_{\min} and a_{\max} .

$$\begin{aligned}\alpha_I &= n_d \sum_{i=1}^{N_m} \Xi_i \int_{a_{\min}}^{a_{\max}} C_{\text{ext}} n_d(a) da \\ \alpha_Q &= n_d \sum_{i=1}^{N_m} \Xi_i \int_{a_{\min}}^{a_{\max}} C_{\text{ext,pol}} n_d(a) da \\ \kappa_Q &= -n_d \sum_{i=1}^{N_m} \Xi_i \int_{a_{\min}}^{a_{\max}} C_{\text{cir1,pol}} n_d(a) da \\ J_I &= n_d \sum_{i=1}^{N_m} \Xi_i \int_{a_{\min}}^{a_{\max}} C_{\text{abs}} n_d(a) da \\ J_Q &= n_d \sum_{i=1}^{N_m} \Xi_i \int_{a_{\min}}^{a_{\max}} C_{\text{abs,pol}} n_d(a) da\end{aligned}$$

For the grain size distributions taken as one of the following:

$$\begin{aligned}n_d(a) &= a^{p_0} \times \exp \left\{ - \left(\frac{a - p_1}{p_2} \right)^{p_3} \right\} && \text{Power-law + Exponential decay} \\ n_d(a) &= \exp \left\{ -0.5 \left(\frac{\log(a) - \log(p_0)}{p_1} \right)^2 \right\} && \text{Log-normal distribution} \\ n_d(a) &= a^{p_0} && \text{Power-law}\end{aligned}$$

The quantity Ξ is the dust material fraction and N_m is the amount of dust grain material where $\sum_{i=1}^{N_m} \Xi_i = 1$.

9 RT Simulation Model

The development and deployment of a robust POLARIS model depends highly on the numerical and physical properties that will govern the simulations. Models are defined as classes containing member functions that are used to describe various components in the simulations.

9.1 Grid Properties

The numerical components become important as they essentially dictate the behavior of all physical quantities and their respective evolutions within the model. Given that the magnetic field model is derived in cylindrical coordinates, it becomes trivial to define the member functions in cylindrical coordinates as well. For the purposes of running the calculations, POLARIS will automatically convert these quantities into cartesian coordinates. A schematic figure of a cylindrical grid embedded in a cartesian cube is demonstrated in figure 23. The generation of

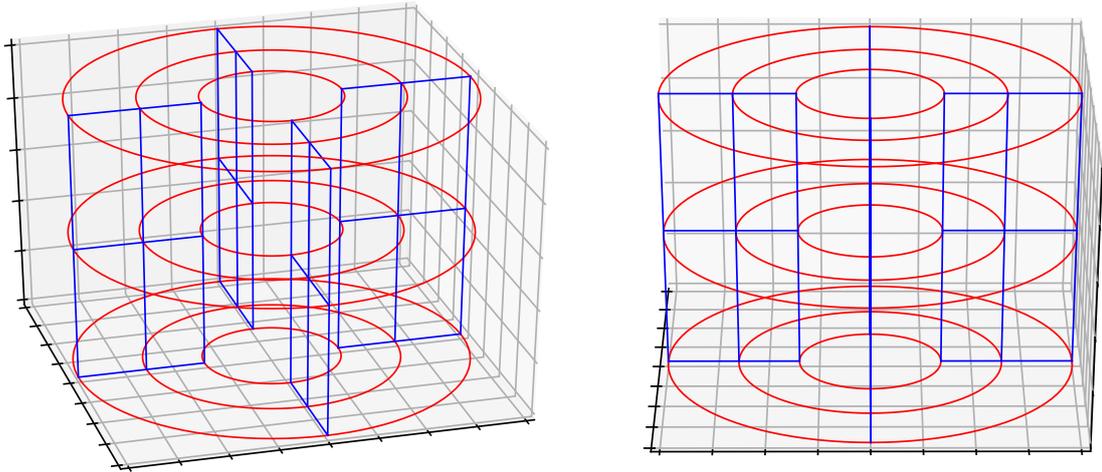


Figure 23: Visualization of the geometry of cylindrical grid generation.

grids are governed by 9 quantities R_{\min} , R_{\max} , Z_{\max} , N_r , N_φ , N_z , f_r , f_φ , f_z where R_{\min} and R_{\max} are the inner and outer radius of the cylinder and the z -axis extends from $-Z_{\max}$ to Z_{\max} . The functions f_r , f_φ and f_z describe the distribution of cells along a given direction in (r, φ, z) . The cell distributions in each direction are given as

$$r_i = \begin{cases} i \frac{R_{\max} - R_{\min}}{N_r} & f_r < 0 \\ R_{\min} + (R_{\max} - R_{\min}) \sin\left(\frac{i\pi}{N_r}\right) & f_r = 1 \\ R_{\max} + \frac{(f_r^i - 1)(R_{\max} - R_{\min})}{f_r^{N_r} - 1} & f_r > 1 \end{cases}$$

$$z_i = \begin{cases} 2i \frac{Z_{\max}}{N_z} & f_z < 0 \\ \frac{\pi}{2} \sin\left(\frac{i\pi}{N_z - 1}\right) & i < \frac{N_z}{2} \text{ and } f_z = 1 \\ \frac{\pi}{2} - \frac{\frac{\pi}{2} (f_z^{N_z/2-i} - 1)}{f_z^{N_z/2} - 1} & i < \frac{N_z}{2} \text{ and } f_z > 1 \\ \frac{\pi}{2} \left(1 - \sin\left(\frac{i\pi}{N_z - 1}\right)\right) & i \geq \frac{N_z}{2} \text{ and } f_z = 1 \\ \frac{\pi}{2} + \frac{\frac{\pi}{2} (f_z^{N_z/2-i} - 1)}{f_z^{N_z/2} - 1} & i \geq \frac{N_z}{2} \text{ and } f_z > 1 \end{cases}$$

and $\varphi_i = 2\pi i/N_\varphi$ for $f_\varphi \neq 0$. All POLARIS simulations will take 50 cells along each cylindrical direction (r, φ, z).

9.2 Physical Properties

The physical components describe the hyper parameters and model the dependencies in the evolutionary properties of the simulation. That is, what factors dictate how quantities of interest behave. The quantities are chosen in a fashion as to best replicate the environment observed surrounding FeSt 1-457 (see [4]).

- **Distance:**

The distance from the system to observer is taken to be 130 pc [4].

- **Dust Composition:**

The dust composition is taken as an oblate shaped grain as it facilitates grain alignment visualization. The composition is 62.5% silicate and 37.5% graphite.

- **Radiation source:**

The grain alignment by radiative torques requires an anisotropic radiation source (see [28]). Therefore, the radiation source implemented is a weakly radiating protostar of temperature $T = 300$ K and radius $R = 4R_\odot$ as the radiation source for the simulations, mimicking a possible observationally undetected first hydrostatic core (see [29]).

Table 9.1: Table of the main model properties.

Main Properties	
Grid Type	Cylindrical
Cells (r)	50
Cells (φ)	50
Cells (z)	50
Distance (pc)	130
Silicate Dust (%)	62.5
Graphite Dust (%)	37.5

9.3 Density Profile

Many prestellar cores have been shown to be accurately modelled by Bonnor-Ebert spheres [21]. That is, the cores are said to be isothermal gas spheres in hydrostatic equilibrium. Taking a spherical gas cloud in hydrostatic equilibrium under the effect of gravity is given as:

$$(9.1) \quad \frac{dp}{dr} = -\frac{Gm(r)\rho(r)}{r^2}$$

where the gas is considered to be in the ideal case and so the equation of state is given as:

$$(9.2) \quad p = c_s^2 \rho(r)$$

where c_s is the isothermal sound speed. Differentiating equation 9.2 with respect to r gives:

$$(9.3) \quad \frac{dp}{dr} = c_s^2 \frac{d\rho(r)}{dr}$$

and so equating equations 9.1 and 9.3 give

$$(9.4) \quad -\frac{Gm(r)\rho(r)}{r^2} = c_s^2 \frac{d\rho(r)}{dr}$$

Where the differential form of the enclosed mass $m(r)$ is given as:

$$(9.5) \quad \frac{dm(r)}{dr} = 4\pi r^2 \rho(r)$$

Taking the derivative of equation 9.4 with respect to r and substituting in equation 9.5 gives the differential equation:

$$(9.6) \quad \frac{1}{r^2} \frac{d}{dr} \left(\frac{r^2}{\rho(r)} \frac{d\rho(r)}{dr} \right) = -\frac{4\pi G \rho(r)}{c_s^2}$$

which has the solution:

$$(9.7) \quad \rho(r) = \frac{c_s^2}{2\pi G r^2}$$

In order to account for the central density $\rho_c = m_n n_c$ [4] (for $n_c = 3.5 \times 10^5 \text{ cm}^{-3}$ being the central number density and $m_n = 2.3m_{\text{H}}$ being the neutral mass), one can instead use a modified isothermal sphere with gas density:

$$(9.8) \quad \rho(r) = \frac{\rho_c}{1 + (r/a)^2}$$

where the radial scale length a is comparable to the Jeans length and is taken to be:

$$a^2 = \frac{A}{2\pi G \rho_c} c_s^2$$

The gas mass is taken to be $M_{\text{gas}} \sim M_{\text{core}} = 3.55M_{\odot}$ [22] and the gas temperature was held constant at 9.5 K [4]. Furthermore, the factor A is taken to be $A = 2$ in order to achieve

a density profile that has a gas mass $M_{\text{gas}} \approx 3.6M_{\odot}$ within the cloud radius $R_{\text{gas}} \sim 0.1$ pc in agreement to the fit performed by Kandori et al. [21]. The core was shown to have little turbulent motion [4] and so no velocity field was applied. In addition to this, the scattering is also taken to be described by the Henyey-Greenstein function. All remaining model properties are summarized in the table below:

Table 9.2: Table of the additional model properties.

Additional Properties	
Gas Mass	$3.55M_{\odot}$
Gas Temperature	9.5 K
Dust Temperature	9.5 K
Velocity Field (v_x, v_y, v_z)	$(0, 0, 0)$
Scattering	Anisotropic
Phase function	Henyey-Greenstein
Grain size distribution	Power-law distribution
Dust to Gas mass ratio	0.01

10 RT Simulation Output

The POLARIS simulation outputs are demonstrated for all three of the developed models thus far. That is, the simulations are run for the full plane model as well as both the left and right plane models where due to axisymmetry, the left and right plane fits have field lines that are mirrored in their respective opposite planes. The expectation is that the orientation of the dust grains will be perpendicular to the local magnetic field taken as an integral quantity through the core which accounts for the integrated scattering and emission properties of the dust grains as well as their respective radiative transfer. The robustness of the model can be back tested and cross validated by comparing the magnetic field lines to the direction of dust grains due to submillimeter emission. In turn this will ultimately lead to the comparison of both the polarimetry data that is observed in FeSt 1-457 (that is the polarization due to near infrared absorption) to the simulated polarimetry data that is output by POLARIS.

10.1 Dust temperature distribution

In order to simulate the dust thermal emission for the grains within the core, one must run a simulation for the temperature distribution. This requires defining a profile for the gas density to which a Bonnor-Ebert-like density is selected. POLARIS performs the simulation by first computing the emissivity given by:

$$(10.1) \quad \varepsilon_l(T_d) = n_d \int C_{\text{abs},\lambda} B_\lambda(T_d) d\lambda$$

This quantity is pre-calculated for all temperatures $T_d \in [1, 4000]$ K. Photons are all given an initial energy per unit of time at some wavelength λ . For each cell in the three dimensional simulation cube having an individual volume given as V_{cell} , the energy per unit time is stored as:

$$(10.2) \quad \dot{E} = n_d \frac{\dot{E}_0}{V_{\text{cell}}} \sum_i C_{\text{abs},\lambda_i} l_i + \Delta \dot{E}$$

for all photon packages along their respective paths l_i throughout the Monte Carlo process. The quantity $\dot{E}_0 \equiv E_0/\Delta t$ is the initial energy per unit of time and $\Delta \dot{E}$ is the offset in the dust heating when there is a pre-defined dust temperature in the model. If the photon interacts within a cell through absorption and re-emission then POLARIS samples a new wavelength from:

$$p = \frac{\int_{\lambda_{\text{min}}}^{\lambda_i} C_{\text{abs},\lambda} D_T[B_\lambda(T_{d,i})] d\lambda}{\int_{\lambda_{\text{min}}}^{\lambda_{\text{max}}} C_{\text{abs},\lambda} D_T[B_\lambda(T_{d,i})] d\lambda}$$

in order to maintain thermal equilibrium. Once all \dot{E} have been computed, a comparison is made with the pre-calculated $\varepsilon_l(T_d)$ values that have been interpolated (POLARIS utilizes cubic splines). The dust temperature distribution is then found by solving

$$(10.3) \quad \varepsilon_l(T_d) = \frac{\dot{E}}{4\pi V_{\text{cell}}}$$

which is valid for systems in local thermodynamic equilibrium.

10.2 Dust thermal emission with grain alignment

Once a dust temperature distribution has been simulated, POLARIS can perform the required dust thermal emission calculations that takes the radiation source in the infrared spectrum ($\lambda = 850\mu\text{m}$). The grain size distribution was taken to follow the MRN power-law distribution where the dust grain number density n_d satisfies

$$\frac{dn_d}{da} \propto n_d^{-3.5}$$

for a minimum grain size $a_{\text{min}} = 0.005\mu\text{m}$ to a maximum grain size $a_{\text{max}} = 2\mu\text{m}$. The synthetic polarization maps are given in figure 25 and demonstrate the polarimetry output of the model viewed face on at $\theta = 0$ and directly head on at $\theta = \pi/2$. An important result of the simulations was that the efficiency of polarization was realized as being as high as 20% which is not consistent with measured quantities that are typically in the 4 – 5% range. This discrepancy is no cause for concern as it is merely an outcome of the constraints and assumptions that went into the POLARIS model. That is, in order to directly infer the morphology of the local magnetic field, perfect alignment of the dust grains are enforced resulting in an increased efficiency in polarization. POLARIS also assumes minimal configurations of dust grains in the simulation where as one typically observes many configurations populating a core in practice. These assumptions do not interfere with the purpose of the analysis which is to utilize radiative transfer simulations to produce polarization maps for the purpose of validating and verifying the robustness of the magnetic field models. Figures 24 and 25 verify that indeed the three dimensional integral magnetic field output from the radiative transfer calculations demonstrate a morphology that is axisymmetric along $\hat{\varphi}$. Viewed at $\theta = \pi/2$ the polarization becomes minimal at the center and viewed at $\theta = 0$ the polarization becomes maximal at the center and drops off radially. These properties are similar to the simulated polarization maps in figure 6 of Kandori et al 2018 [22] (see orientations at 0 degrees and 90 degrees) where they demonstrate a distribution that is spherical and maximal at the center. In figure 24, the synthetic polarization maps are demonstrated for the full plane fit adopting $R = 0.1$ pc and the newly fitted model for $R = 0.2$ pc representative of the core at an earlier contraction state. Furthermore, figure 25 shows the polarization maps representative of the fits to the left and right planes of the data. Additionally, model performance can be crudely assessed by overlaying the polarization segments of FeSt 1-457 with the synthetic maps of figures 24 and 25. In figure 26, the red vectors are representative of the near infrared measurements of the polarization segments of FeSt 1-457 and the black vectors are the simulated outputs from POLARIS rotated by 90 degrees to demonstrate the implied morphology of the integrated magnetic field through the core. In these figures one should seek to have the red and black vectors parallel to one another.

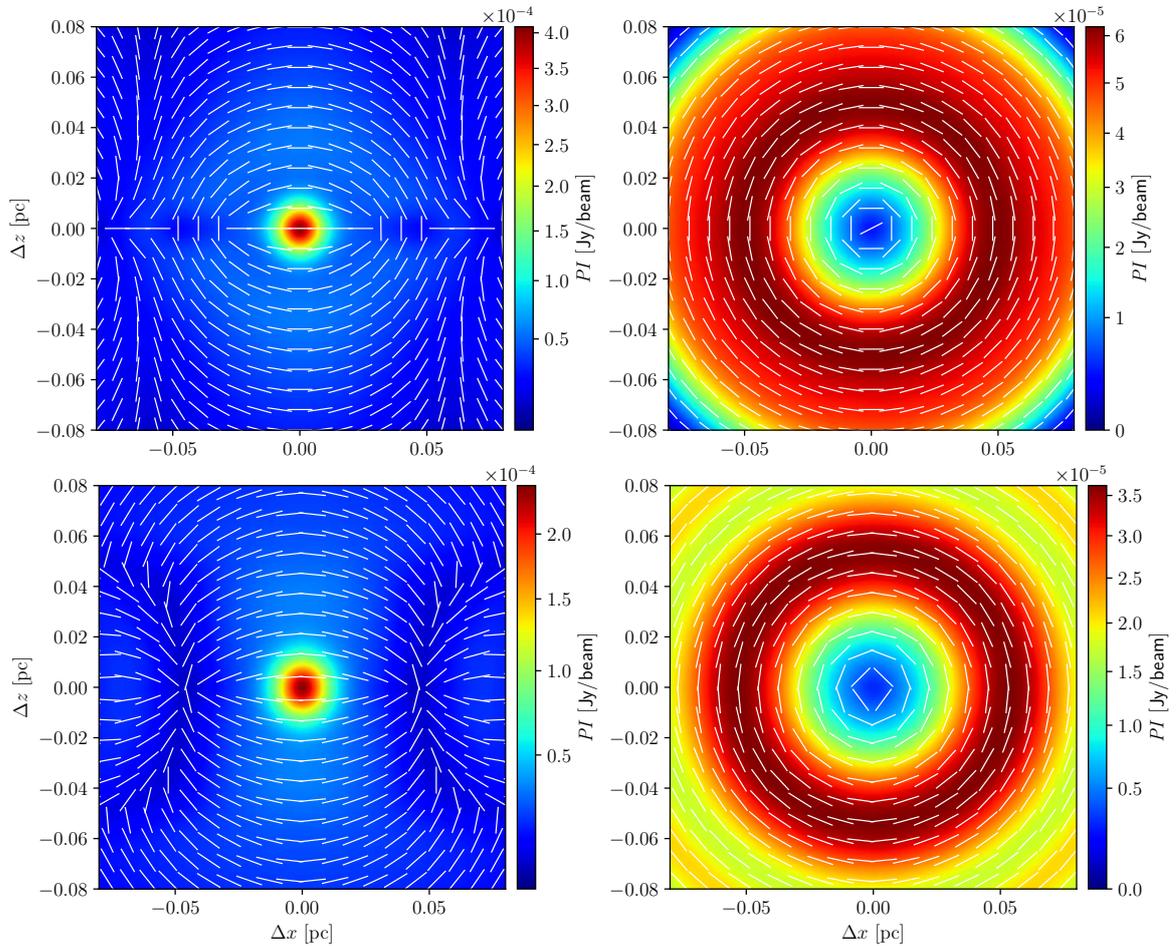


Figure 24: Synthetic polarization maps given in terms of the polarization intensity (in Jansky per beam). Grain alignment oriented perpendicular to the magnetic field lines. **Left:** Oriented face on to the observer $\theta = 0$. **Right:** Oriented top down to the observer $\theta = \pi/2$. **Top to bottom:** Full plane fit model adopting $R = 0.1$ pc and the newly fitted model adopting $R = 0.2$ pc.

10.3 Mass-to-Flux ratio

This section focuses on studying the contraction mechanism for the prestellar core. This is done by calculating the mass to flux ratio for the cloud. Recall that the critical mass to flux ratio is given as:

$$\left(\frac{M}{\Phi}\right)_{\text{crit}} = \frac{1}{2\pi\sqrt{G}}$$

and so the dimensionless parameter μ is defined as:

$$(10.4) \quad \mu \equiv \frac{(M/\Phi)}{(M/\Phi)_{\text{crit}}} = 2\pi\sqrt{G} \left(\frac{M}{\Phi}\right)$$

which will be used to study the magnetic properties of the core. The flux is taken for the magnetic field threading the equatorial plane given as:

$$(10.5) \quad \Phi = \int_A \mathbf{B} \cdot d\mathbf{A} = 2\pi \int_0^R B_z(r, 0) \times r dr$$

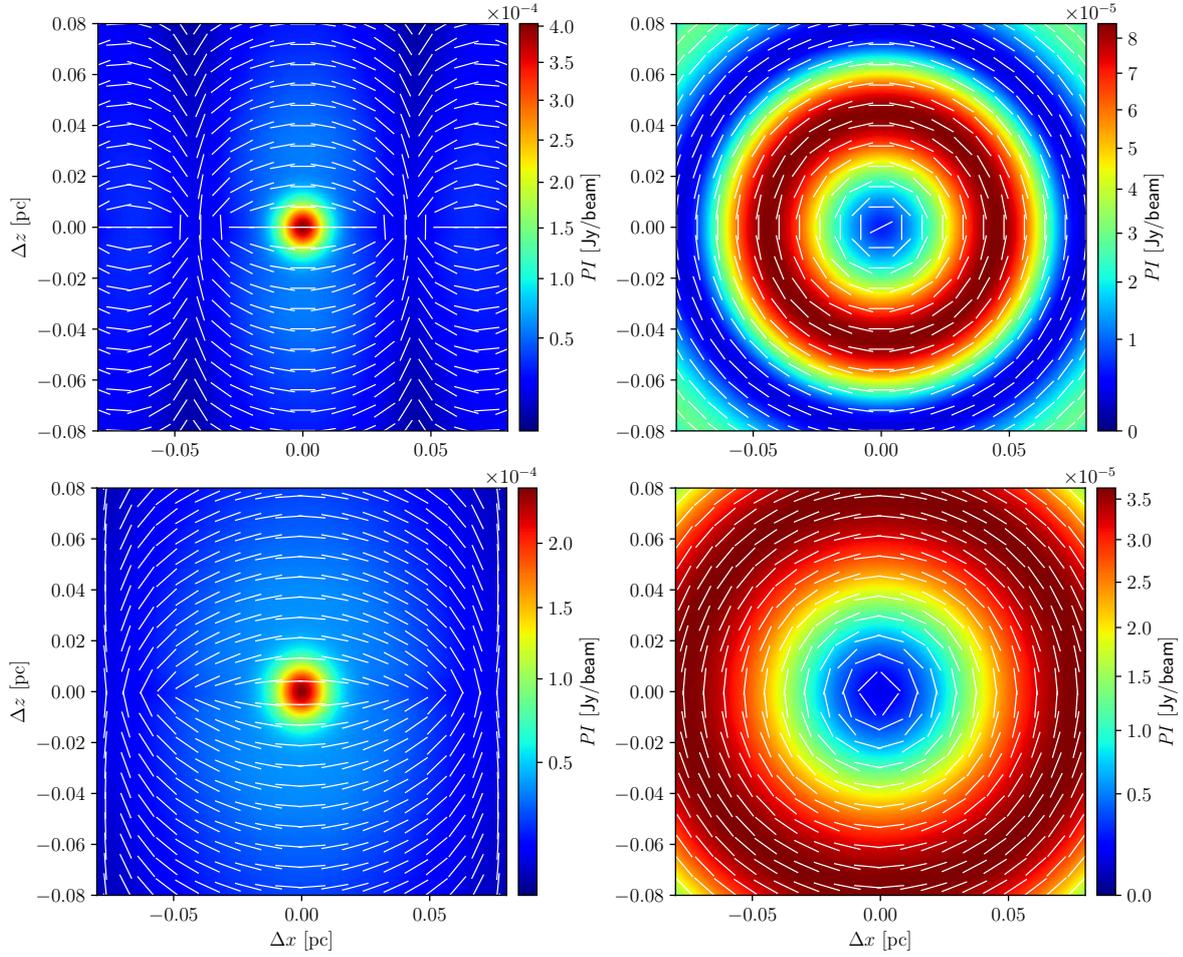


Figure 25: Synthetic polarization maps given in terms of the polarization intensity (in Jansky per beam). Grain alignment oriented perpendicular to the magnetic field lines. **Left:** Oriented face on to the observer $\theta = 0$. **Right:** Oriented top down to the observer $\theta = \pi/2$. **Top to bottom:** Left plane fit and the right plane fit.

Inserting the form of the model magnetic field the dimensionless flux $\hat{\Phi}$ is given:

$$(10.6) \quad \hat{\Phi} = 2\pi \int_0^1 \left(\sum_{m=1}^3 \beta_m J_0(a_{m,1}\tilde{r}) \left[\operatorname{erfc}\left(\frac{a_{m,1}\eta}{2}\right) + \operatorname{erfc}\left(\frac{a_{m,1}\eta}{2}\right) \right] + 1 \right) \times \tilde{r} d\tilde{r}$$

The integral portion of equation 10.6 can be found in analytic form for limits of integration $[0, x]$

$$(10.7) \quad 2\pi \int_0^x \tilde{B}_z(\tilde{r}, 0) \times \tilde{r} d\tilde{r} = \pi x^2 \left(\frac{4}{x} \sum_{m=1}^3 \frac{\beta_m}{a_{m,1}} J_1(a_{m,1}x) \operatorname{erfc}\left(\frac{a_{m,1}\eta}{2}\right) + 1 \right)$$

Inserting for $x = 1$ gives:

$$(10.8) \quad \hat{\Phi} = \pi \left(4 \sum_{m=1}^3 \frac{\beta_m}{a_{m,1}} J_1(a_{m,1}) \operatorname{erfc}\left(\frac{a_{m,1}\eta}{2}\right) + 1 \right)$$

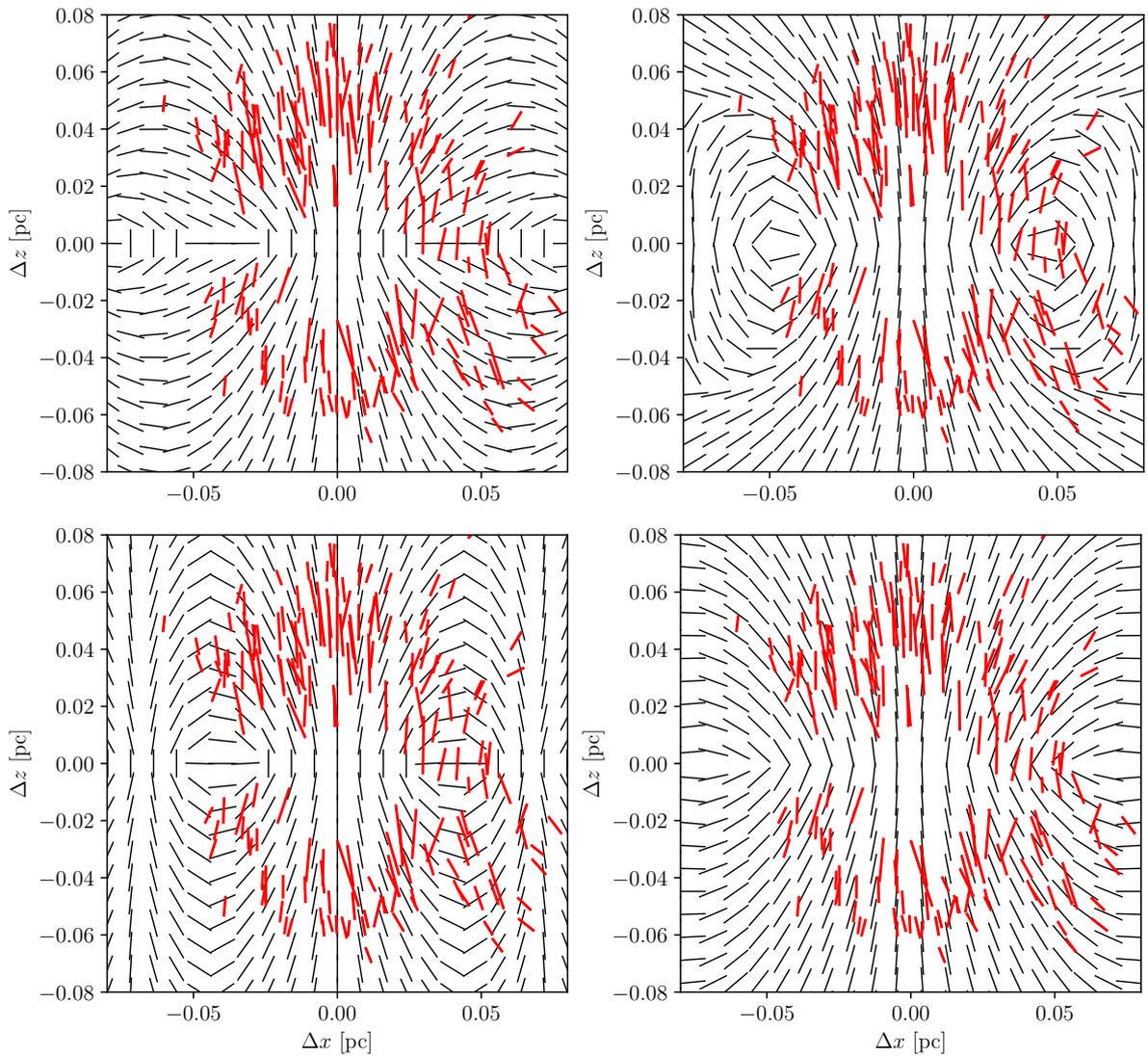


Figure 26: Synthetic polarization maps with magnetic field vectors in black. The polarization segments are oriented by 90 degrees to demonstrate the implied morphology of the integrated magnetic field through the core. The superposition of the polarimetry data (red lines) are demonstrated with the inferred magnetic field (black lines). **Top left to right:** Full plane fit model adopting $R = 0.1$ pc and the newly fitted model adopting $R = 0.2$ pc. **Bottom left to right:** Left plane fit and the right plane fit.

which evaluates to $\hat{\Phi} = \pi$ and therefore the flux of the core's generated magnetic field can be given as:

$$(10.9) \quad \Phi_{\text{core}} = \pi R^2 B_0$$

This shows that the model is fitting a magnetic field structure that is within a region $r = R$ and assumes the cloud condensed from the background field within the same region under flux freezing. The dimensionless parameter μ for the full plane fit can be written as:

$$(10.10) \quad \mu = 0.8 \left(\frac{M_{\text{core}}}{3.6 M_{\odot}} \right) \left(\frac{R}{0.1 \text{ pc}} \right)^{-2} \left(\frac{B_0}{15 \mu\text{G}} \right)^{-1}$$

For the adopted core mass and radius this means that the core is somewhat subcritical ($\mu \lesssim 1$) if the magnetic field is taken as $B_0 = 15 \mu\text{G}$, the strength for the intercore medium of molecular clouds using the OH Zeeman effect [30]. Alternatively taking $B_0 = 5 \mu\text{G}$, representative of the general interstellar medium [31] gives $\mu \approx 2$; a core that is somewhat supercritical where the contraction is mainly gravitationally driven. This is comparable with the parameter found in Kandori et al 2018 [22] and infers that the magnetic forces within the core are not strong enough to prevent dynamical collapse. Furthermore, the mass to flux ratio for the model adopting $R = 0.2 \text{ pc}$ can be estimated within the half-radius (0.1 pc) by setting the upper limit of integration in equation 10.6 to $\tilde{r} = 0.5$. The integral now evaluates the normalized flux to be $\hat{\Phi} = 0.44$ for $r \leq 0.5R = 0.1 \text{ pc}$. The mass to flux ratio is therefore given as:

$$(10.11) \quad \mu = 4.6 \left(\frac{M_{\text{core}}}{3.6 M_{\odot}} \right) \left(\frac{R}{0.2 \text{ pc}} \right)^{-2} \left(\frac{B_0}{15 \mu\text{G}} \right)^{-1}$$

which yields a decisively supercritical core. Additionally, the mass to flux ratio would still be supercritical even if the background field B_0 is taken as $5 \mu\text{G}$ extending to $R = 0.2 \text{ pc}$. The flux-freezing model of Myers et al. [18] was applied to FeSt 1-457 and demonstrated a fit to the inner region $r \approx R_{\text{gas}}/2$ that appears to be consistent with observations. Outside this region, the model magnetic field lines revert to vertical morphology along the z -axis. Given the strong degree of curvature at $r \gtrsim 0.04 \text{ pc}$, the flux-freezing model has difficulty resolving these due to the model reverting to straight field lines. This can be interpreted as a flux-freezing contraction starting at a relatively dense state at a gas radius $R_{\text{gas}} \lesssim 0.1 \text{ pc}$ adopted from the gas column density profile. The outer region corresponding to highly curved polarization segments could also be indicative of an approximately flux-freezing contraction. Given that the model developed fits the magnetic field independent of the density profile, flux freezing is not required at an earlier contraction state. The full plane fit model adopting $R \approx R_{\text{gas}} = 0.1 \text{ pc}$ fits polarization segments (including outer ones) relatively well, but results in a subcritical mass to flux ratio for a background field $B_0 = 15 \mu\text{G}$ [22]. Instead for the fit fixing $R = 0.2 \text{ pc}$, the density within the relevant region ($r = 0.1 \text{ pc}$) follows a profile similar to a Bonnor-Ebert sphere. Adopting the same intercore medium $B_0 = 15 \mu\text{G}$, the inner region now becomes supercritical with outer

region likely being subcritical. This can be accomplished by rapid ambipolar diffusion that can be modelled by large-scale nonlinear flow [17, 32–37]. In that case, the nonlinear flow is responsible for bringing together ambient subcritical interstellar gas giving rise to the highly curved field lines at the core’s outer edge. Strong magnetic field gradient in the inner dense region along with low ionization gives rise to rapid ambipolar diffusion. Once the inner structure becomes supercritical, the dense core undergoes collapse into star formation. Furthermore, a quantitative assessment is made by computing a mean residual estimate of the relative angles each polarization vector makes with the z -axis for the magnetic field directions inferred from the POLARIS output along with the those from the best-fit midplane magnetic field models. Given that POLARIS outputs more polarization segments than that from the data set of FeSt 1-457, the calculations are performed only incorporating the closest dust grain to each polarization segment. The dust grains are rotated by 90 degrees to get the inferred magnetic field direction. For the midplane model, the relative angle is calculated using the relation

$$(10.12) \quad \varphi = \tan^{-1} \left(\frac{\sqrt{B_x^2 + B_y^2}}{B_z} \right) \sim \tan^{-1} \frac{B_r}{B_z}$$

evaluated at the midpoints of the observed polarimetry data. For the POLARIS output, the mean residuals are

$$\overline{\Delta}_\varphi^{(F)} = 0.4081 \text{ rad} \quad \overline{\Delta}_\varphi^{(L)} = 0.4508 \text{ rad} \quad \overline{\Delta}_\varphi^{(R)} = 0.2941 \text{ rad}$$

which are found to be in good agreement with the midplane model residuals

$$\overline{\Delta}_\varphi^{(F)} = 0.4064 \text{ rad} \quad \overline{\Delta}_\varphi^{(L)} = 0.5260 \text{ rad} \quad \overline{\Delta}_\varphi^{(R)} = 0.3061 \text{ rad}$$

where the superscripts 'F', 'L' and 'R' denote the full, left and right plane fit models respectively. This agreement gives reassurance that the midplane magnetic field model can be used to find good fits to the observed polarization segments. To assess the goodness of fit further, the standard assumption is made that the model residuals are normally distributed. This hypothesis is tested utilizing Pearson’s chi-squared test and a residual whiteness test. Pearson’s chi-squared test defines the statistic

$$(10.13) \quad \chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

where O_i and E_i are representative of the observed and expected counts respectively. In an effort to compensate for any potential measurement errors or systematic errors in the data extraction process. In doing so, outliers have been removed utilizing Matlab’s `rmoutliers` function and the resulting chi-squared estimate yields a statistic of 6.51. This result does not reject the null hypothesis of the model residuals being normally distributed and the resulting histogram is demonstrated in figure 27. Furthermore, a residual whiteness test is performed

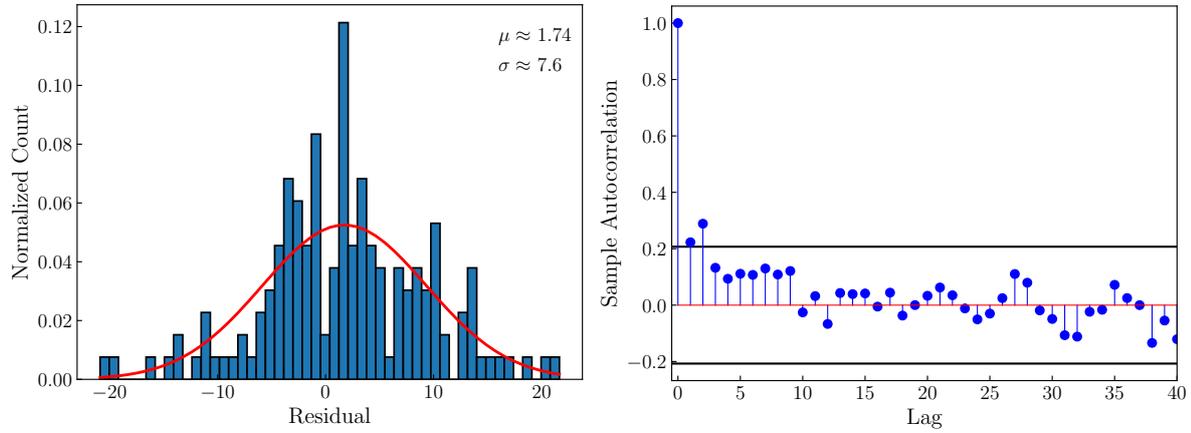


Figure 27: Statistical figures assessing the model's goodness of fit. **Left:** Histogram of residuals fit to a normal distribution with mean $\mu \approx 1.74$ and standard deviation $\sigma \approx 7.6$. **Right:** Residual autocorrelation function up to 40 lags with 99% confidence interval.

using the residual autocorrelation function testing the autocorrelation between residuals at different lags. Adopting a 99% confidence interval, the right panel of figure 27 shows that the majority ($\sim 95\%$) of lags fall within the confidence interval (ignoring lag $k = 0$ being representative of the autocorrelation of a data point with itself). This is indicative of the model residuals being approximately normally distributed.

11 Incorporating Radial Current Densities

In deriving the magnetic field equations 6.4 & 6.5, the total current density was assumed to be separable:

$$(11.1) \quad j(r, z) = f(r)g(z)$$

where the vertical portion was taken to be Gaussian of the form:

$$(11.2) \quad g(z) = e^{-z^2/h^2}$$

where h is the vertical scale length. From appendix A1.1, the magnetic vector potential is given as:

$$(11.3) \quad A(r, z) = \frac{2}{h\sqrt{\pi}} \sum_{m=1}^{\infty} k_m e^{-h^2\lambda_m/4} J_1(\sqrt{\lambda_m}r) \int_{-\infty}^{\infty} g(\eta) e^{-\sqrt{\lambda_m}|z-\eta|} d\eta$$

where taking $g(\eta) = e^{-\eta^2/h^2}$ allowed the integral to be evaluated analytically. The radial current density is buried in the parameters k_m :

$$(11.4) \quad k_m = \frac{2h\pi^{3/2}e^{h^2\lambda_m/4}}{cR^2\sqrt{\lambda_m} [J_2(\sqrt{\lambda_m}R)]^2} \int_0^R f(\xi) J_1(\sqrt{\lambda_m}\xi) \xi d\xi$$

Thus far, the model has been implemented by taking these parameters as part of a fit rather than evaluating them directly. This methodology did not allow one to define a radial current as was done with $g(z)$. By explicitly quantifying $f(r)$, an alternative approach can be introduced to modelling magnetic field using equations 6.4 & 6.5. Take the radial current to follow a similar Gaussian profile:

$$(11.5) \quad f(r) = e^{-r^2/l^2}$$

where l is taken as a radial characteristic scale length. Equation 11.4 now becomes

$$(11.6) \quad k_m = \frac{2h\pi^{3/2}e^{h^2\lambda_m/4}}{cR^2\sqrt{\lambda_m} [J_2(\sqrt{\lambda_m}R)]^2} \int_0^R e^{-\xi^2/l^2} J_1(\sqrt{\lambda_m}\xi) \xi d\xi$$

and will be evaluated numerically using the trapezoidal method. The magnetic field patterns are generated by varying the ratio of respective scale lengths h/l . The series is again truncated at $m = 3$ terms and the ratio is increased in accordance to $h/l \in [0.01, 0.1, 0.75]$. Defining $\mathbf{k} = [k_1, k_2, k_3]^T$ to be the vector of coefficients yields:

$$(11.7) \quad \mathbf{k}^{(1)} = \begin{bmatrix} 3.29 \\ 1.61 \\ 0.68 \end{bmatrix} \times 10^{-9}, \quad \mathbf{k}^{(2)} = \begin{bmatrix} 3.32 \\ 1.66 \\ 0.72 \end{bmatrix} \times 10^{-8},$$

$$\mathbf{k}^{(3)} = \begin{bmatrix} 4.14 \\ 6.81 \\ 19.52 \end{bmatrix} \times 10^{-7},$$

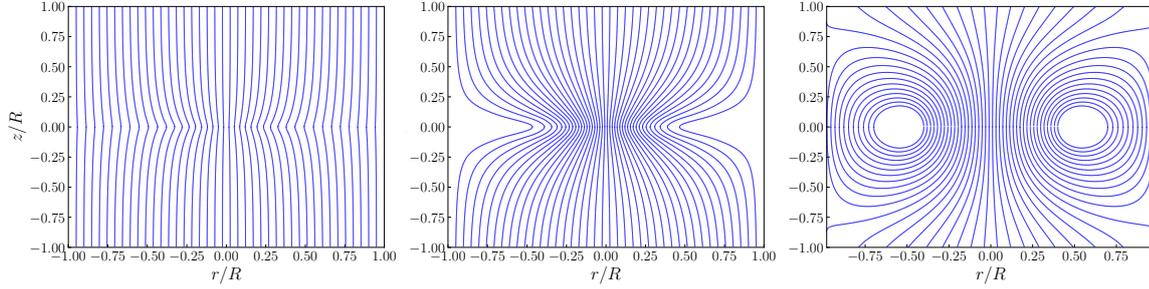


Figure 28: Magnetic field morphology demonstrated for the scale length ratio $h/l \in [0.01, 0.1, 0.75]$ (from left to right) with respective $B_c/B_0 \in [2.63, 13.69, 51.95]$.

where the superscripts correspond to $(h/l_1, h/l_2, h/l_3) = (0.01, 0.1, 0.75)$. The respective magnetic field morphologies are given in figure 28 which are similarly taken to demonstrate a midplane cut. Interestingly, as the ratio h/l is increased, the magnetic field becomes more pinched. The effect of varying the scale length along any given direction will make the current more concentrated for smaller scale lengths and more disperse for large scale lengths. This observation can give insight that highly pinched morphologies observed in practice may not always be an outcome of small B_0 , but can also arise from the relative distributions of current along the r - and z -directions. This result provides an alternative methodology to the analysis of hourglass magnetic field. Directly evaluating the integral in equation 11.4 provides a potential extension in modelling magnetic field. For an appropriate selection of $f(r)$, the total number of parameters can be reduced from $m + 3$ down to a maximum of 4 (R, B_0, h, l). Furthermore, there are no restrictions to the selection of $f(r)$ and $g(z)$. In future work, a survey of various profiles will be considered and combinations of such can be implemented to studying how the magnetic field is affected for different pairs of $f(r)$ and $g(z)$.

12 Conclusion

In comparing a polarimetry map of the prestellar core FeSt 1-457 with a single plane of the model magnetic field and a simulated synthetic polarization map, many estimates governing the core's properties can be inferred. Visually, there appears to be good agreement between the model magnetic field and the simulated synthetic dust emission polarization map obtained from POLARIS. The fits suggest that FeSt 1-457 is in either a transcritical or in a mild supercritical state. The highly curved field lines in the magnetic field morphology suggest that the core has condensed from some larger distance $r > R$. The study performed demonstrated the capabilities of an analytic magnetic field model used in fitting the ratio B_r/B_z along the midplane to the direction of polarization segments. For an independently estimated background field B_0 , the model calculates the magnetic field strength at all points within the model space. The model adopts the physically plausible Gaussian function along the vertical direction, which can be extended to survey more general distributions in the future which can include its radial counterpart.

SMALL SCALE STRUCTURE – PROTOSTELLAR CORES

Collaborator: Masahiro Machida, Kyushu University

Key Words: protostars: young stellar objects – magnetic fields: MHD – polarization

13 Introduction

The magnetic field governing protostellar cores plays an important role in star formation. The gas dynamics are influenced by the magnetic field, which regulates the gravity and can dictate the rate of formation [38, 39]. Polarimetry measurements provide the best tool for studying and inferring the magnetic properties. Inferring magnetic field in protostellar cores can be challenging due to the limited availability of high resolution polarization maps. Although observationally measured polarization maps of have been sparse and difficult to come by, there are still some that have been made available in recent years [40–42]. Given these limitations, theoretical models are commonly used in interpreting these measurements and inferring the respective observable properties. The evolutionary properties governing protostellar cores can be modelled through magnetohydrodynamic (MHD) simulation. Due to the light ionization of the clouds [13], numerical MHD simulations must be non-ideal to account for the magnetic field not being fully coupled to the plasma. These simulation outputs form the basis for a model that can be incorporated with observed polarimetry data to compute synthetic polarization maps. These synthetic maps are used to infer the magnetic field at a given orientation that is integrated along the observer’s line of sight. By numerically solving the resistive MHD equations over a nested grid, a series of protostellar models can be constructed at various length scales, resolving smaller and smaller regions. For each model output, the radiative transfer code POLARIS can be used to simulate a respective polarization map at a given length scale. The MHD simulations are used to produce high resolution models within a simulation cube having side lengths as small as $\approx 90 - 180$ au. These length scales can be smaller than much of the observed polarimetry data currently available and enable predictions for future higher resolution observations.

14 Non-Ideal Magnetohydrodynamic Simulation

In order to simulate the dense, star-forming core of molecular clouds, non-ideal magnetohydrodynamic simulations are used as the basis for the model. This is a direct extension to the model discussed in the previous chapter. In these dense regions, the clouds are lightly ionized [13] and because of this the magnetic field cannot be coupled to the plasma and so the flux frozen property is no longer valid. It is at this stage that the plasma needs to be treated in a non-ideal MHD fashion to account for neutrals, ions, electrons and other charged material. The simulations were performed by Machida et al. [43–46] using their 3-D nested MHD code that solves the resistive MHD equations including self-gravity:

$$(14.1) \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

$$(14.2) \quad \rho \frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla p + \mathbf{j} \times \mathbf{B} - \rho \nabla \Phi$$

$$(14.3) \quad \frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B}$$

$$(14.4) \quad \nabla^2 \Phi = 4\pi G \rho$$

for the density ρ , the velocity \mathbf{v} and the magnetic field \mathbf{B} . Here, Φ is the gravitational potential, \mathbf{j} is the electric current density, p is the pressure, η is the resistivity and G is the gravitational constant.

14.1 Initial State

As an initial state, the core adopts a Bonnor-Ebert density profile with a central density of $6 \times 10^5 \text{ cm}^{-3}$ and isothermal temperature $T = 10 \text{ K}$. The initial core mass and radius are taken to be $M_{\text{cl}} = 2M_{\odot}$ and $R_{\text{cl}} = 1.2 \times 10^4 \text{ au}$, respectively. The simulation also enhances the cloud density by a factor $f = 1.68$ to promote the contraction. For the initial core, a uniform magnetic field of $B_0 = 3.2 \times 10^{-5} \text{ G}$ is taken with a rigid rotation of $\Omega_0 = 1.3 \times 10^{-13} \text{ s}^{-1}$ where the rotation axis is parallel to the magnetic field. Additionally, the normalized mass-to-flux ratio in the initial state is taken to be $\mu_0 = 2$ which is normalized to the critical value $1/(2\pi G^{1/2})$. Finally, the ratio of thermal and rotational energy to gravitational energy are $\alpha_0 = 0.42$ and $\beta = 0.024$, respectively.

14.2 Ohmic Dissipation & Sink Treatment

The Ohmic dissipation is modelled by the effective resistivity [47] in the induction equation 14.3. The resistivity η is adopted from Nakano et al. [48] and is quantitatively estimated to being a function of the number density and the temperature:

$$(14.5) \quad \eta = \frac{740}{\chi_e} \sqrt{\frac{T}{10\text{K}}} \left[1 - \tanh \left(\frac{n}{10^{15} \text{ cm}^{-3}} \right) \right] \text{ cm}^2 \text{ s}^{-1}$$

for the ionization degree of the gas

$$\chi_e = 5.7 \times 10^{-4} \left(\frac{n}{\text{cm}^{-3}} \right)^{-1}$$

where T is representative of the gas temperature and n is the number density. Finally, the MHD simulation adopts a sink at the center of the simulation grid to accelerate the calculation. Taking the 'sink radius' $r_{\text{sink}} = 2$ au, the region $r < r_{\text{sink}}$ removes gas having number density $n > 10^{12} \text{ cm}^{-3}$ from the simulation grid and adds it to the protostar in the form of gravity for each timestep. For $n < 10^{12} \text{ cm}^{-3}$ the equation of state is taken to represent an isothermal ideal gas. Further details are given in Machida et al. [49] and this methodology allows for a longer duration of circumstellar disk evolution.

14.3 Mathematical Methods

In order to effectively model the governing evolutionary properties of the core, the resistive MHD solutions of Machida et al. are used [43–46]. Their numerical code solves the 3-D resistive MHD equations on a nested grid using a Monotonic Upstream-centered Scheme for Conservation Laws (MUSCL) [43, 50] to the conservative form of the MHD equations:

$$(14.6) \quad \frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot F(\mathbf{U}) = 0$$

where \mathbf{U} represents the vector of conserved variables and $F(\mathbf{U})$ represents the flux of conserved variables. Figure 29 gives a schematic representation of a nested grid in two dimensions. Imposing this geometry refines the mesh toward the center of the domain and keeps the mesh coarse closer to the boundaries. The numerical scheme scales the simulation on the nested grid up to three dimensions and groups the outputs into levels representative of a given scale within the simulation cube. Each level is made up of 64 cells across each dimension where the quantities of interest are evaluated at the center of each cell located at:

$$x_i^{(l)} = 2^{-l+1} \left(i + \frac{1}{2} \right) h \quad y_j^{(l)} = 2^{-l+1} \left(j + \frac{1}{2} \right) h \quad z_k^{(l)} = 2^{-l+1} \left(k + \frac{1}{2} \right) h$$

for a step size h that is halved at each level and the superscript l being a given simulation level. Each positional index i, j, k is bounded as $-N/2 \leq i, j, k \leq N/2 - 1$. The cube has volume $V = L^3 = (Nh)^3$ and the simulation runs for 13 levels total (only levels 5 - 13[†] are considered since earlier stages do not present significant systematic evolution of the protostellar core). The grid lengths are given in table 14.1 and the reader is referred to appendix A1.4 for details surrounding how the MHD gravitational solver is implemented on the nested grid.

[†]Level 13 is only used in the later stages of the core's evolution. Earlier stages only simulated up to level 12.

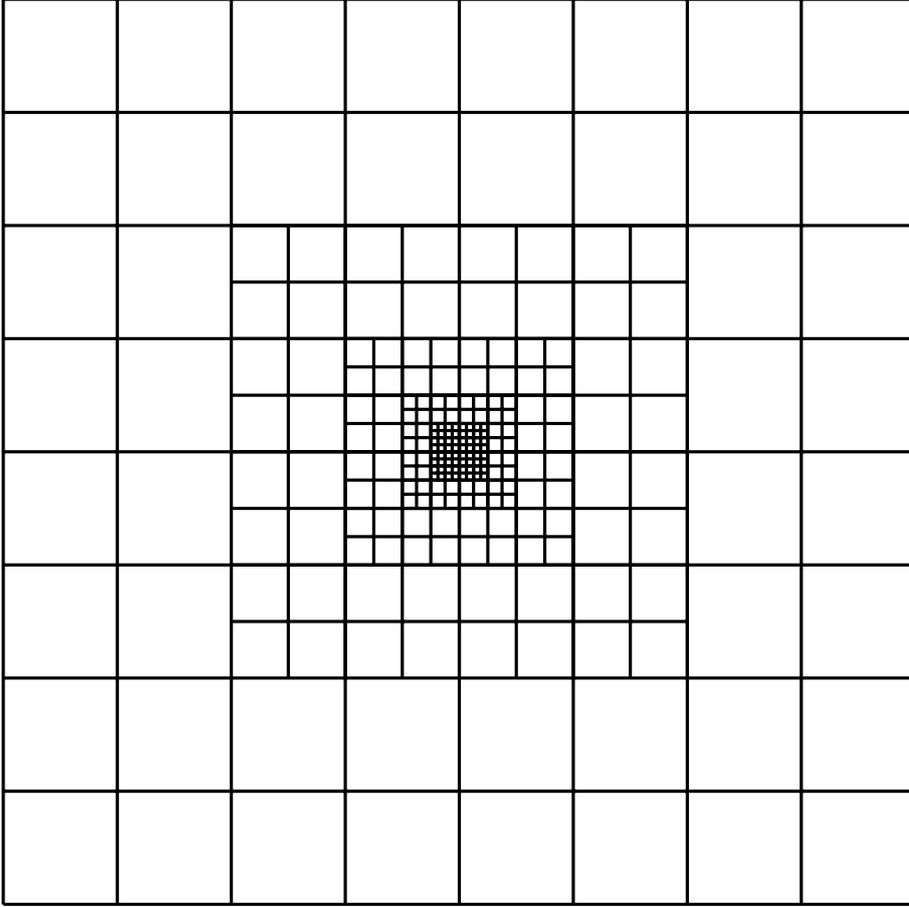


Figure 29: Exemplary geometry of a nested grid demonstrated for a total of 5 levels of refinement.

Table 14.1: Nested Grid lengths.

Level	Box length (L)
$l = 5$	23655.5 au
$l = 6$	11827.8 au
$l = 7$	5913.88 au
$l = 8$	2956.94 au
$l = 9$	1478.47 au
$l = 10$	739.235 au
$l = 11$	369.617 au
$l = 12$	184.809 au
$l = 13$	90.9602 au

14.4 MHD Simulation Output

The MHD simulations run up to $t \approx 89547$ years, which corresponds to a mass of $M \approx 0.3M_{\odot}$ in the core's evolution. The simulation outputs are demonstrated through figures at each level within the nested grid. Therefore, each figure displays 8 or 9 different plots from level 5 up to and including levels 12 and 13. The primary analysis of the simulations studies the core's structure for three stages in the evolutionary process for when $M = 0.01M_{\odot}$, $M = 0.2M_{\odot}$ and $M = 0.28M_{\odot}$. These outputs will be extended to simulating synthetic polarization maps at all levels in the nested grid. This methodology will allow for high resolution figures of the small scale regions within protostellar cores. Furthermore, to better visualize the core's dynamics, figure 30 demonstrates the time evolution of the plasma outflow as the core has evolved to $M = 0.15M_{\odot}$, $M = 0.25M_{\odot}$ and $M = 0.30M_{\odot}$. These plots are taken at level 7 in the nested grid as a midplane cut at the 32nd cell and are shown in the $x - z$, $y - z$ and $x - y$ planes.

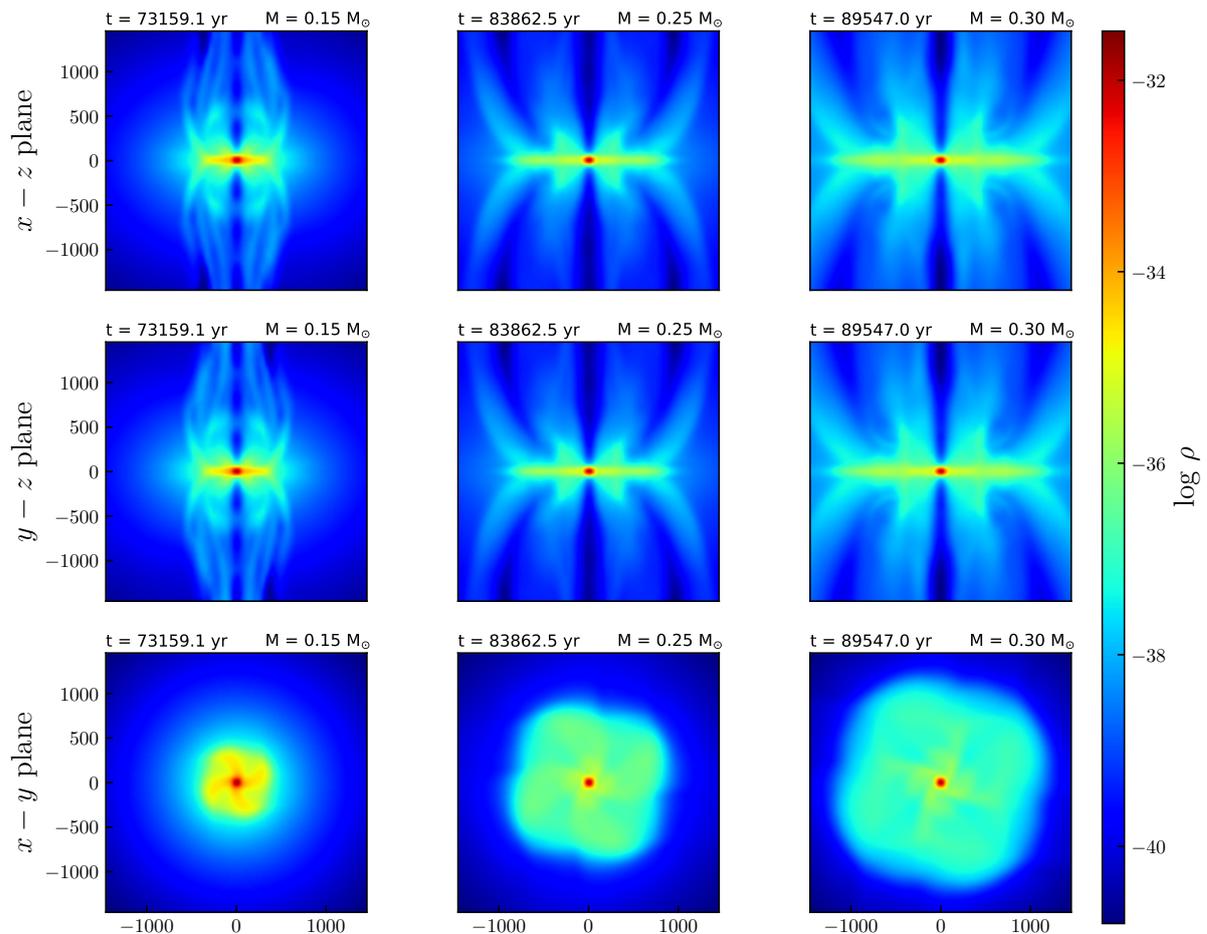


Figure 30: Density (g/cm^3) given to demonstrate the time evolution of the outflow as the core evolves. Figure is given for the level 7 stage in the simulation grid. Axis units are given in au.

Mass: $0.01M_{\odot}$

• Temperature:

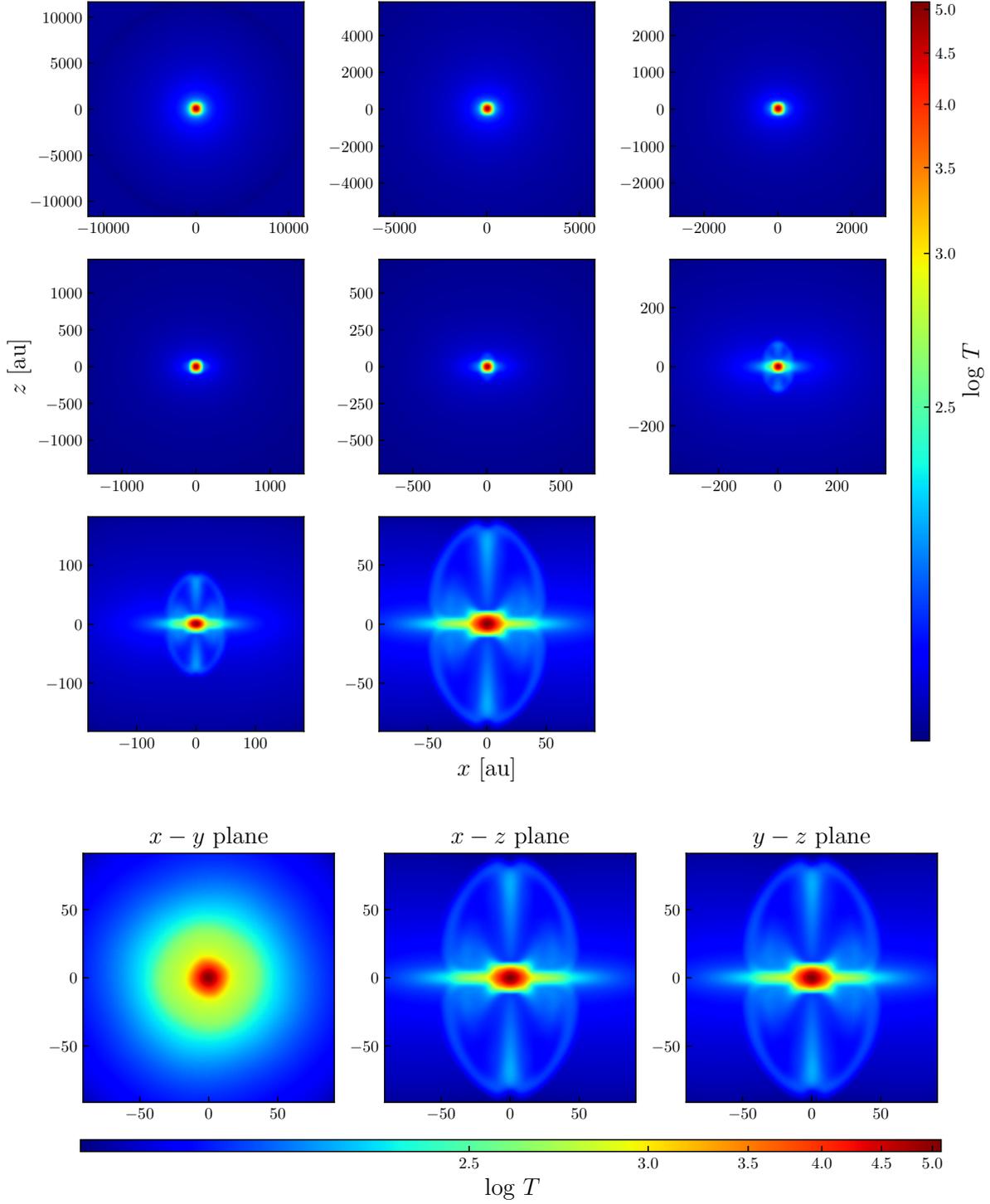


Figure 31: Top: Temperature (K) for all levels taken at a slice through the y plane at the 32nd cell. **Bottom:** Temperature distribution for level 12 at slices through the x , y and z plane at the 32nd cell.

• Density:

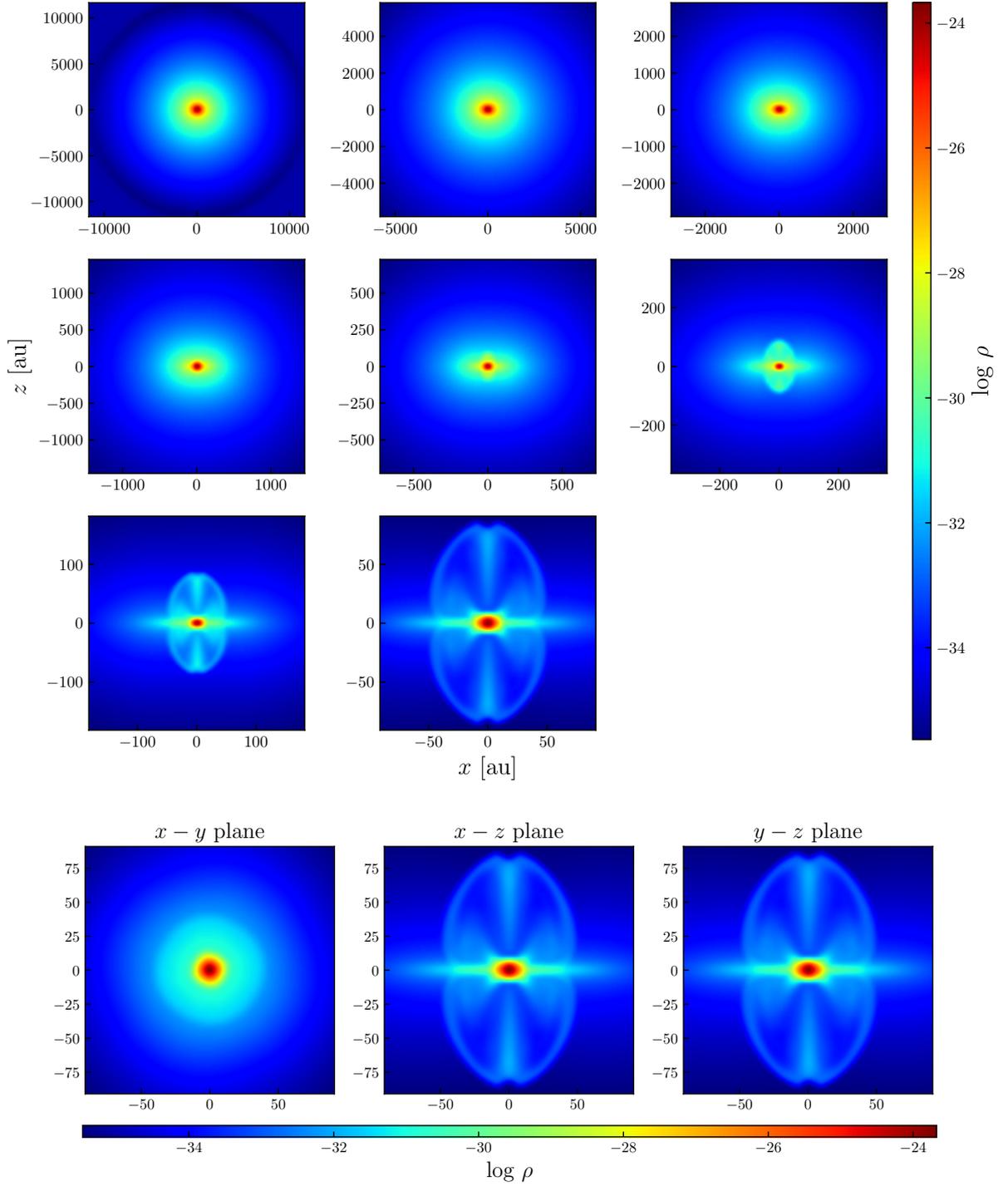


Figure 32: **Top:** Density (g/cm^3) for all levels taken at a slice through the y plane at the 32nd cell. **Bottom:** Density distribution for level 12 at slices through the x , y and z plane at the 32nd cell.

• Magnetic Field:

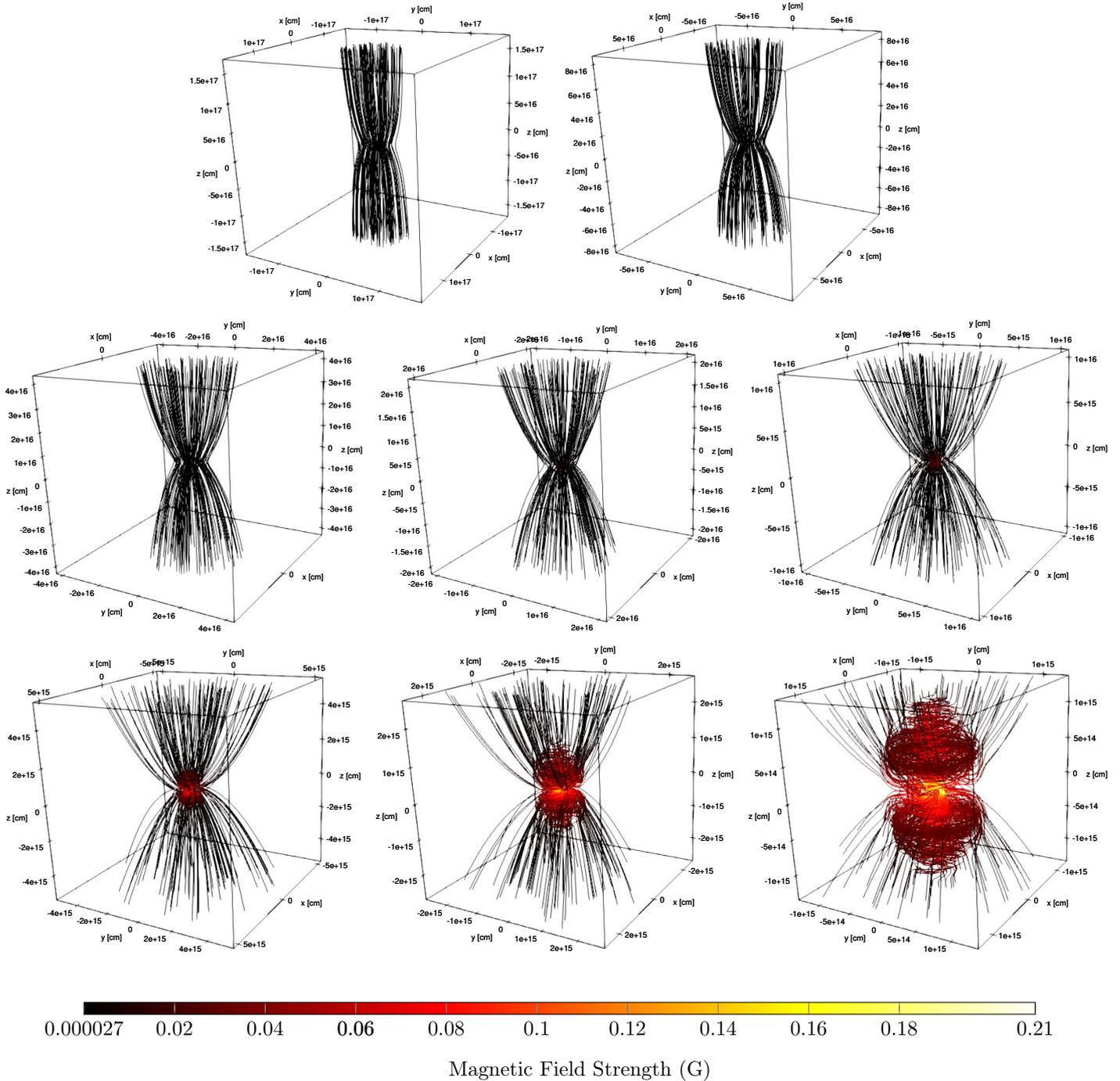


Figure 33: Three dimensional magnetic field lines for levels 5 - 12. **Top:** Levels 5 and 6. **Middle:** Levels 7, 8 and 9. **Bottom:** Levels 10, 11 and 12. The bottom panels corresponding to the higher levels show the increasing influence of toroidal magnetic fields at progressively smaller scales.

• Velocity Field:

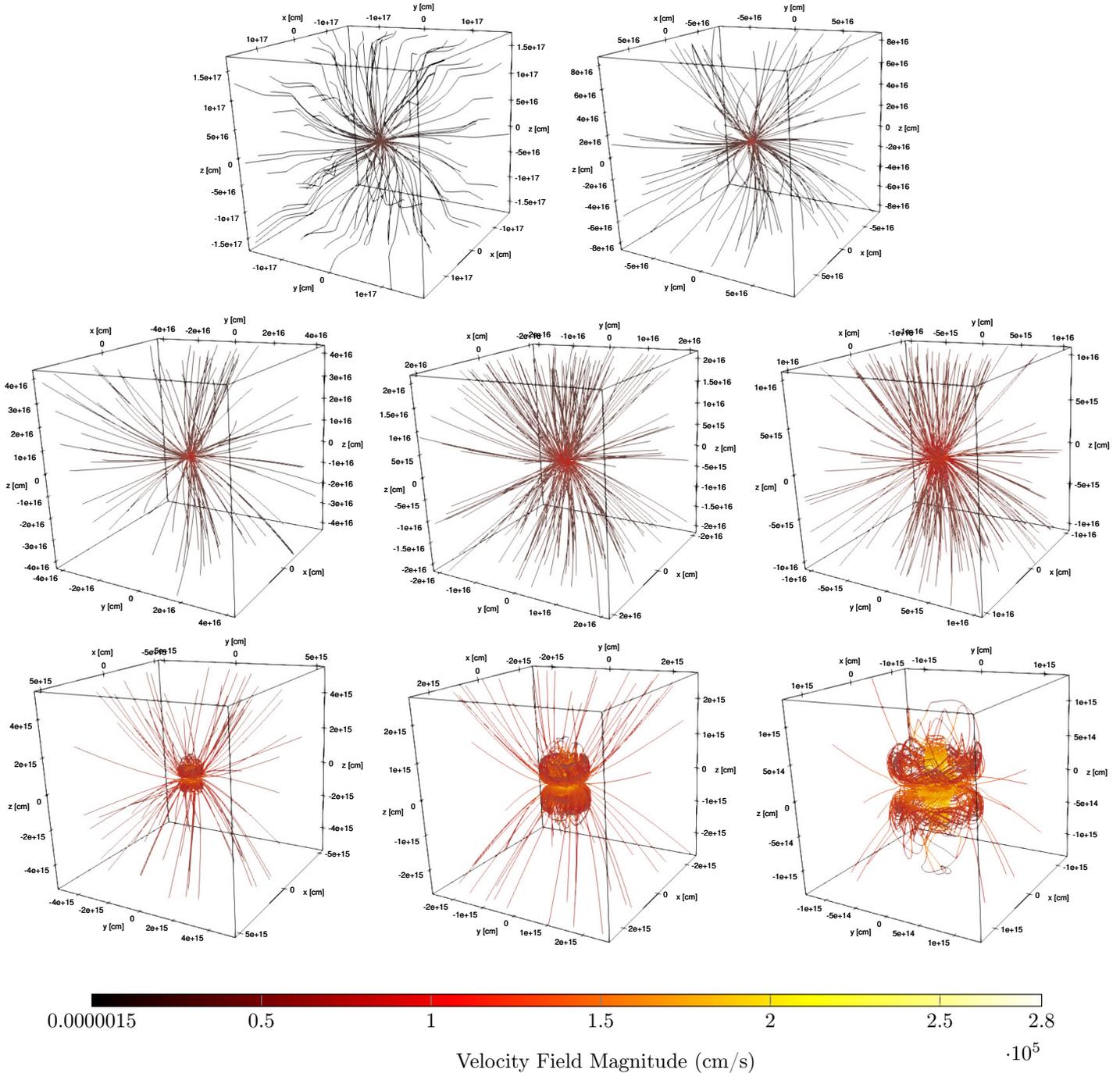


Figure 34: Three dimensional velocity field for levels 5 - 12. **Top:** Levels 5 and 6. **Middle:** Levels 7, 8 and 9. **Bottom:** Levels 10, 11 and 12.

Mass: $0.2M_{\odot}$

• Temperature:

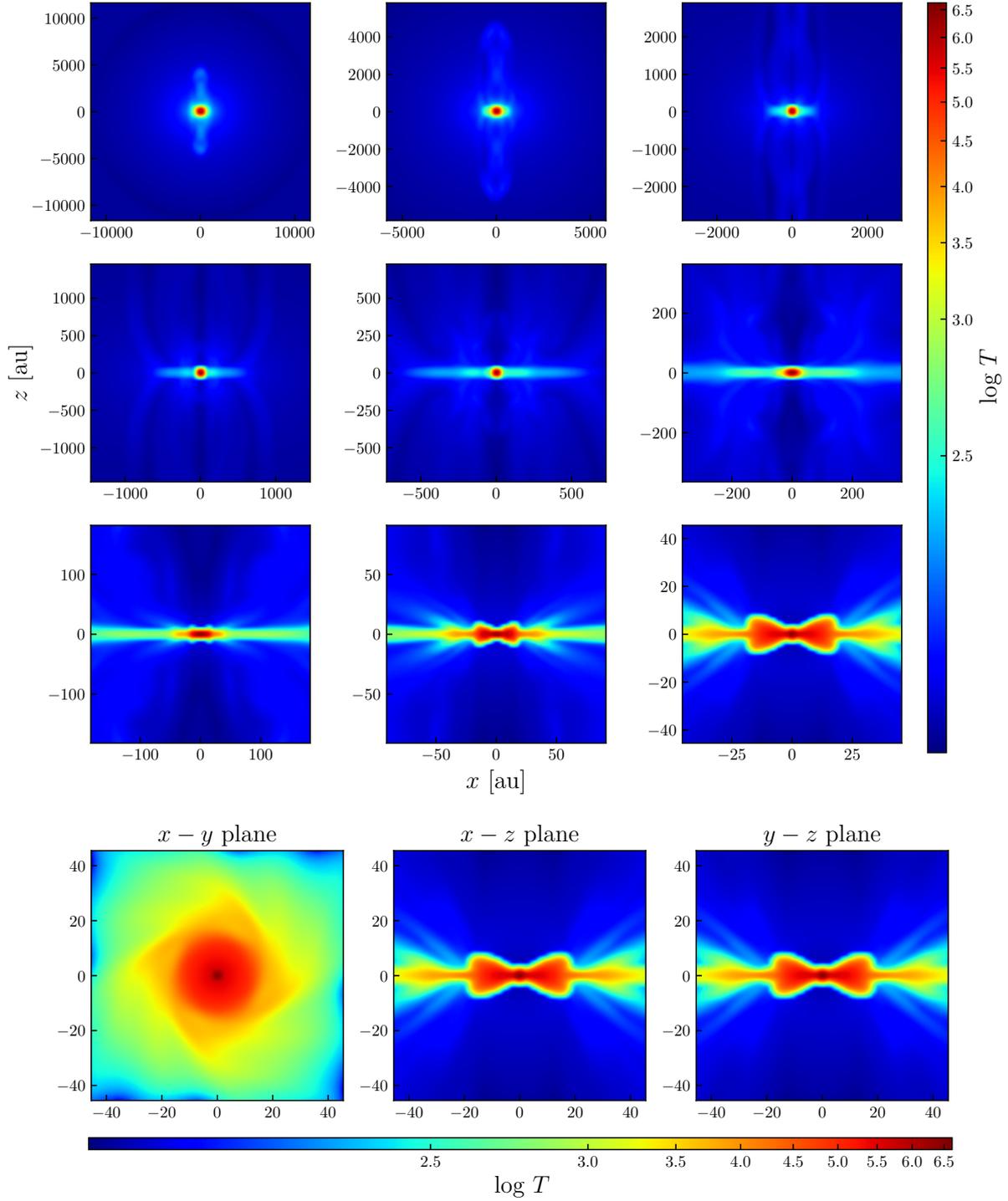


Figure 35: Top: Temperature (K) for all levels taken at a slice through the y plane at the 32nd cell. **Bottom:** Temperature distribution for level 13 at slices through the x , y and z plane at the 32nd cell.

• Density:

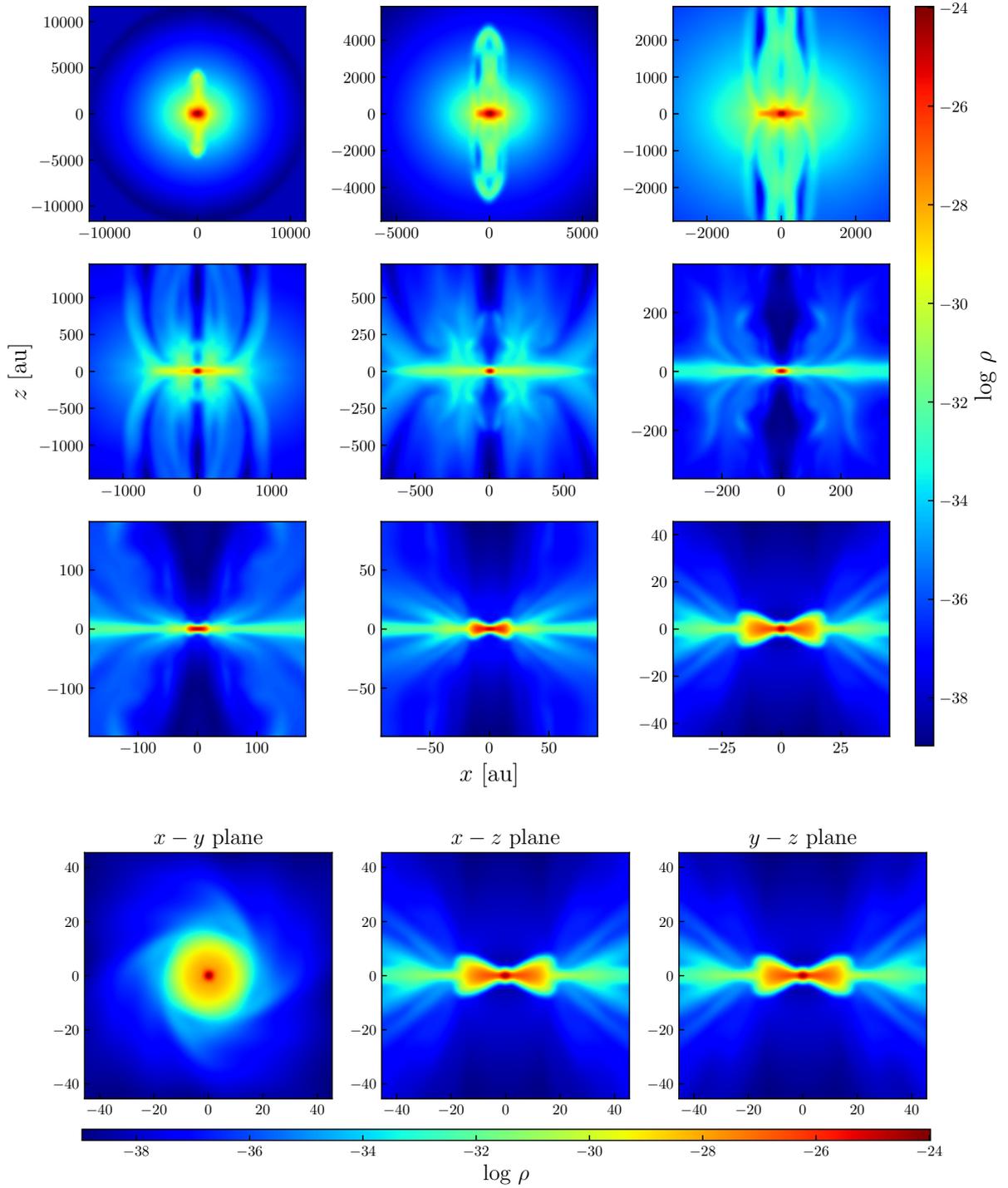


Figure 36: **Top:** Density (g/cm^3) for all levels taken at a slice through the y plane at the 32nd cell. **Bottom:** Density distribution for level 13 at slices through the x , y and z plane at the 32nd cell.

• Magnetic Field:

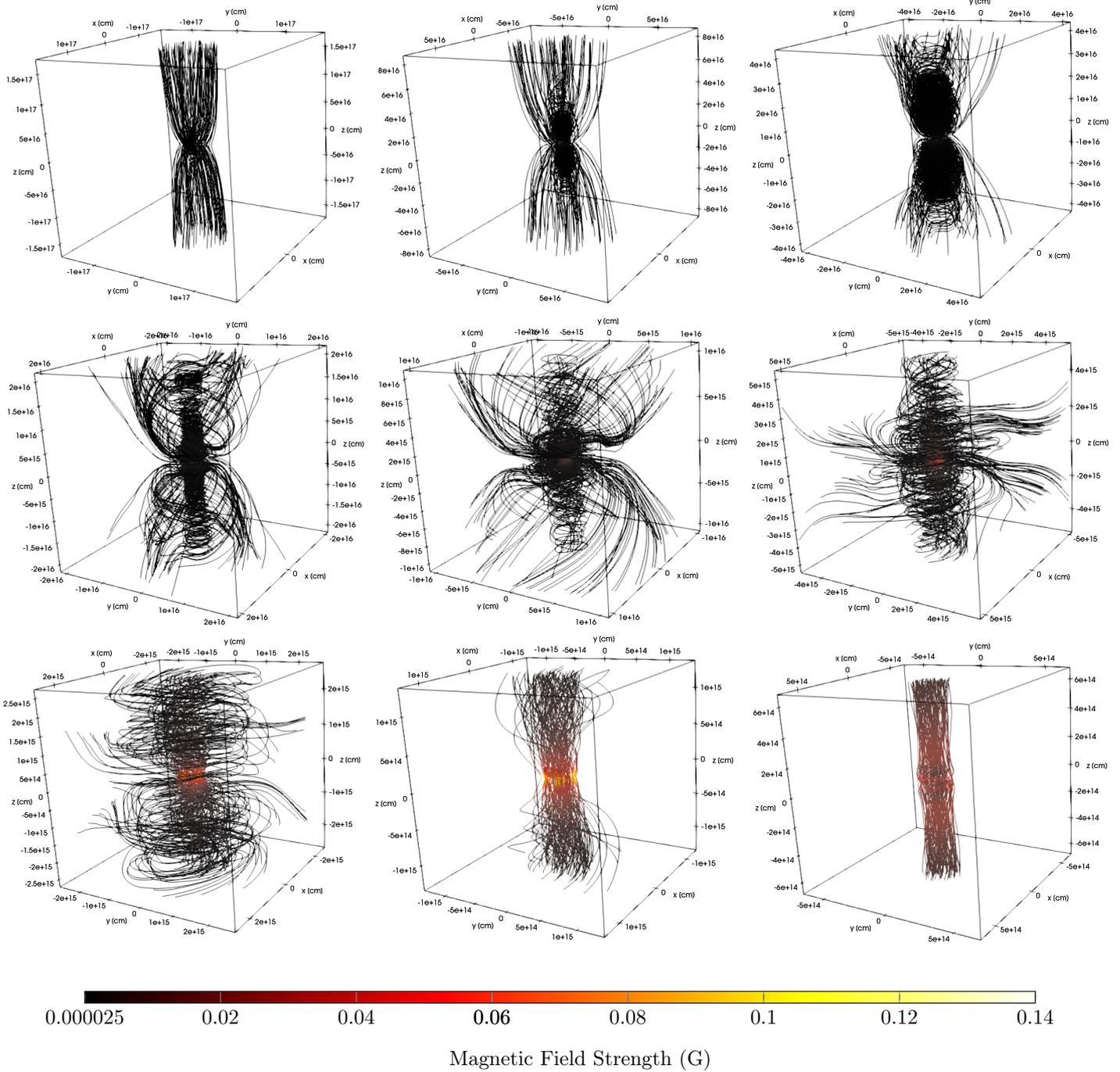


Figure 37: Three dimensional magnetic field lines for levels 5 - 13. **Top:** Levels 5, 6 and 7. **Middle:** Levels 8, 9 and 10. **Bottom:** Levels 11, 12 and 13. The bottom panels corresponding to the higher levels show the increasing influence of toroidal magnetic fields at progressively smaller scales.

• Velocity Field:

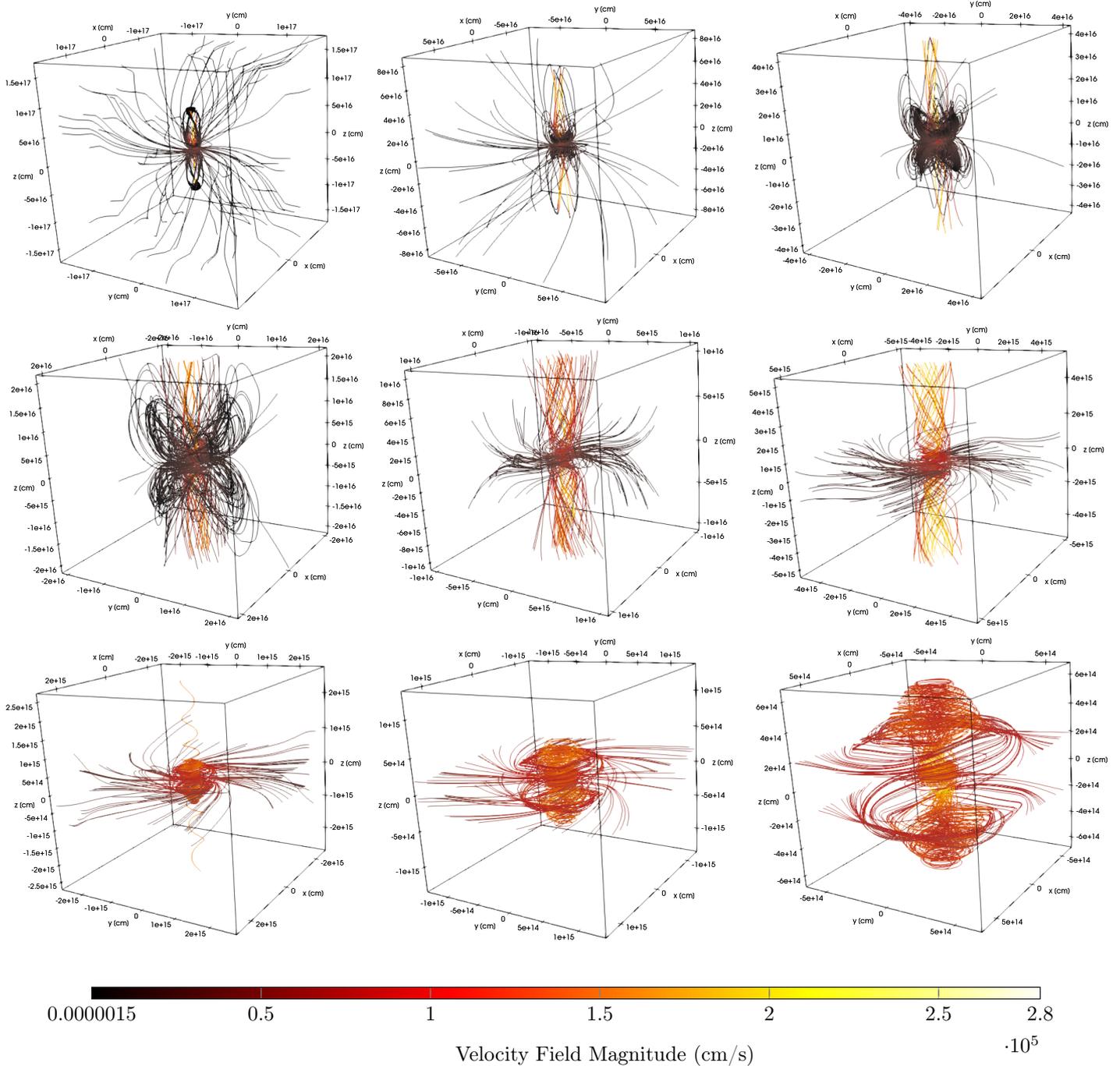


Figure 38: Three dimensional velocity field lines for levels 5 - 13. **Top:** Levels 5, 6 and 7. **Middle:** Levels 8, 9 and 10. **Bottom:** Levels 11, 12 and 13.

Mass: $0.28M_{\odot}$

• **Temperature:**

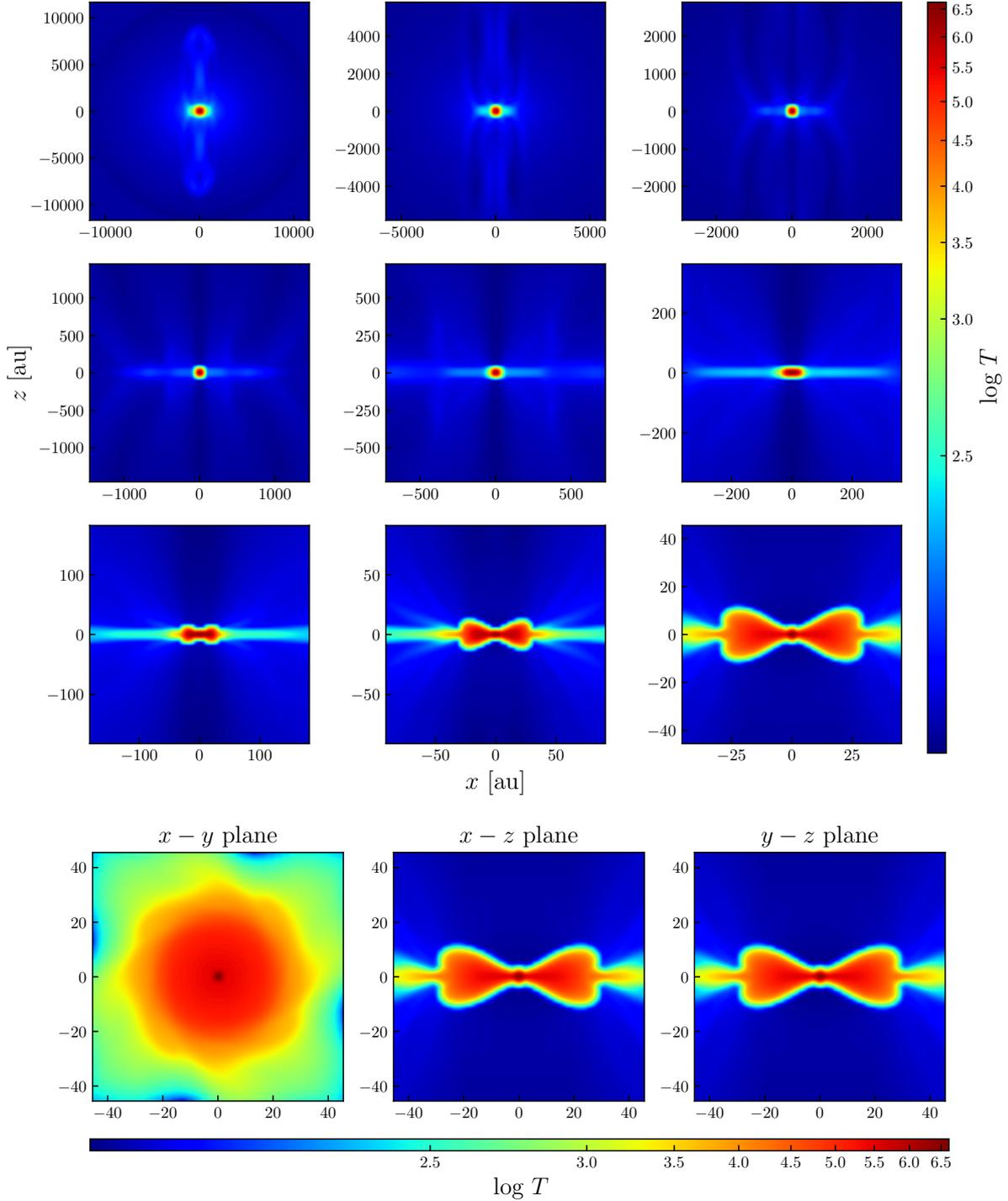


Figure 39: Top: Temperature (K) for all levels taken at a slice through the y plane at the 32nd cell. **Bottom:** Temperature distribution for level 13 at slices through the x , y and z plane at the 32nd cell.

• Density:

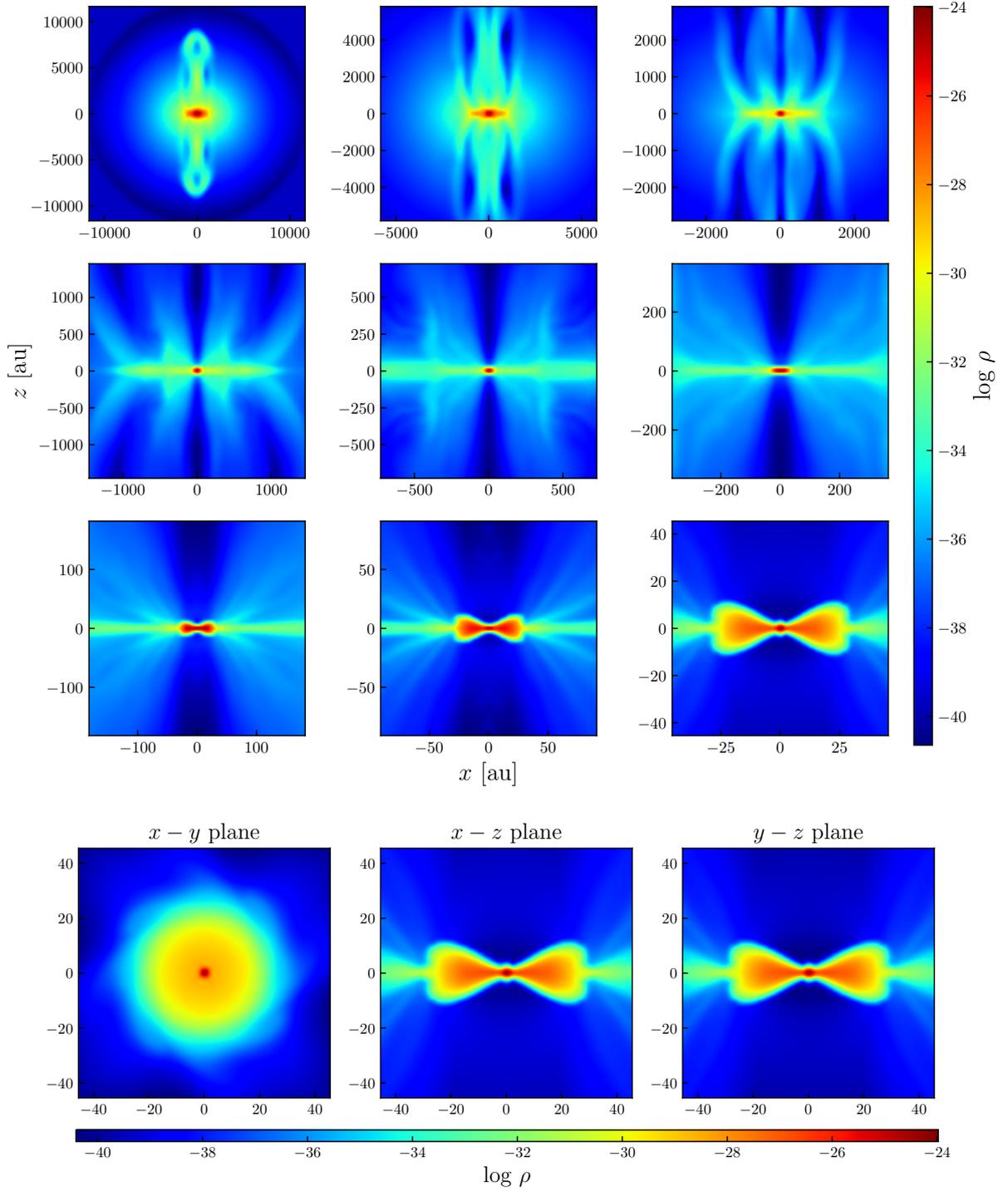


Figure 40: Top: Density (g/cm^3) for all levels taken at a slice through the y plane at the 32nd cell. Bottom: Density distribution for level 13 at slices through the x , y and z plane at the 32nd cell.

14. NON-IDEAL MAGNETOHYDRODYNAMIC SIMULATION

• Magnetic Field:

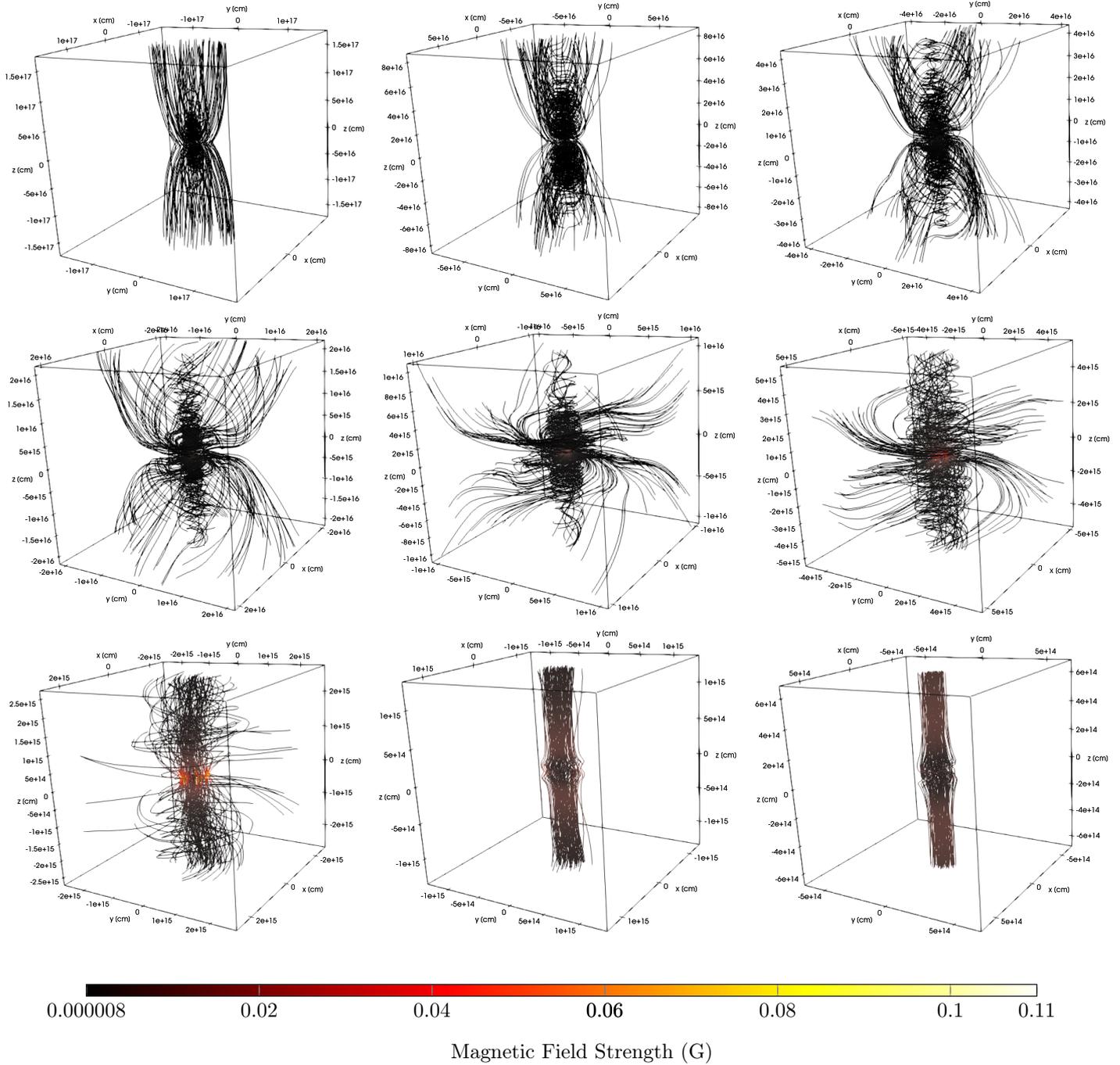


Figure 41: Three dimensional magnetic field lines for levels 5 - 13. **Top:** levels 5, 6 and 7. **Middle:** Levels 8, 9 and 10, **Bottom:** Levels 11, 12 and 13. The bottom panels corresponding to the higher levels show the increasing influence of toroidal magnetic fields at progressively smaller scales.

• Velocity Field:

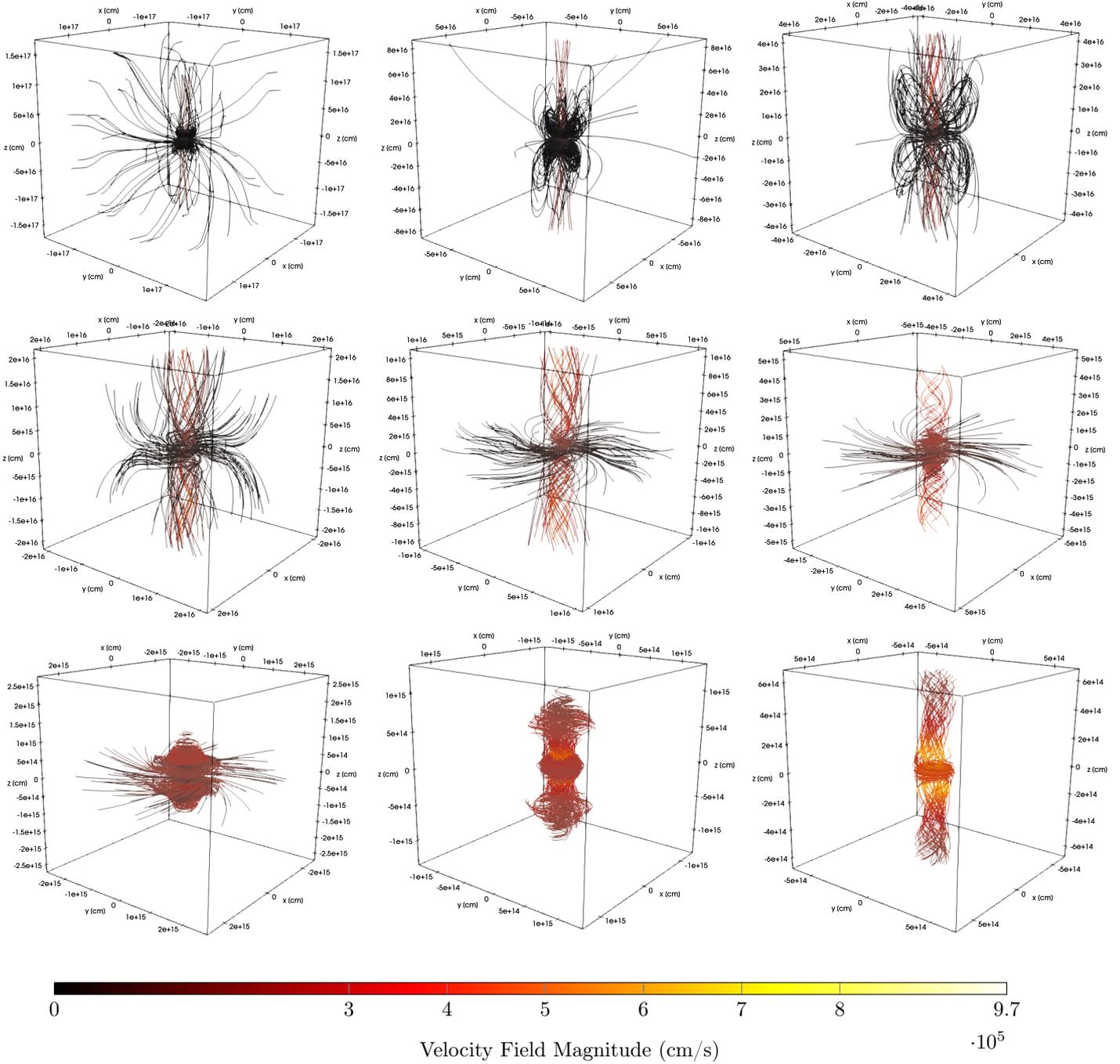


Figure 42: Three dimensional velocity field lines for levels 5 - 13. **Top:** Levels 5, 6 and 7. **Middle:** Levels 8, 9 and 10. **Bottom:** Levels 11, 12 and 13.

15 Numerical MHD in POLARIS

Simulating synthetic polarization maps can be used to extrapolate and infer various properties governing the protostellar core when comparing with observations. By modelling the core's evolutionary properties using the resistive MHD equations, synthetic polarization maps can be made by constructing a radiative transfer calculation from the MHD solution. These MHD outputs will serve as the foundation for the development of the radiative transfer model in POLARIS that will calculate the emergent intensity and polarization of emission from dust grains in the core. With no closed form solution to any of the governing properties, the simulation data is explicitly specified at each cell within the POLARIS simulation grid. That is, the grid index position (i, j, k) is initialized with the values $B_{x,y,z}$, $v_{x,y,z}$, ρ and T . This ensures that at some grid index POLARIS can access the correct data value from the external input in order to perform the radiative transfer calculations.

15.1 Cartesian Grid with Refinement

The radiative transfer simulations in POLARIS for small scale structure are performed on an adaptive octree (cartesian) grid. The mesh is designed such that each cell is declared to be either a leaf or a branch. That is, if the cell is defined to be a leaf it is made up of data and if the cell is defined to be a branch, it is part of a refinement to a subset of leafs. An exemplary diagram is shown in figure 43 for up to 3 levels.

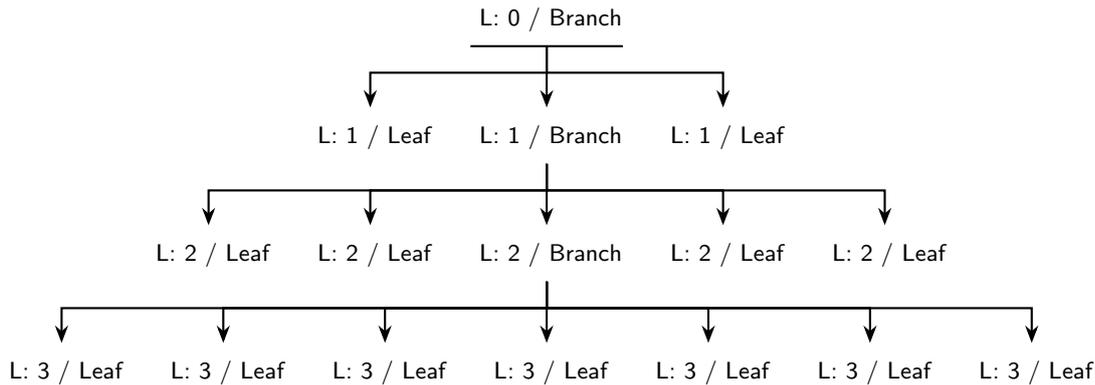


Figure 43: Exemplary tree diagram of an adaptive octree grid.

Each increasing level represents a higher level of refinement in the grid where level 0 represents the entire cube and level 3 in this case represents the maximal refinement level. This refinement technique is similar to the nested grid, but does not limit the refinement to the center of the mesh and instead allows refinement anywhere across the cube.

15.2 Model Properties

The temperature and gas density profiles are all taken from the MHD simulation data along with the velocity and magnetic field. The POLARIS simulations adopt an octree grid with 7 total levels of refinement.

- **Distance:**

The distance from the system to observer is taken to be 140 pc.

- **Dust composition:**

The dust is taken as an oblate shaped grain with composition 62.5% silicate and 37.5% graphite.

- **Radiation source:**

The radiation source for the POLARIS and MHD simulation were similarly taken to be a T-Tauri star. T-Tauri [36] stars comprise the youngest class of stars with a mass $\lesssim 2M_{\odot}$ and is located at the center of the simulation cube. Furthermore, the mass is taken to match that of its respective MHD counterpart. The remaining properties are summarized in table 15.1.

Table 15.1: Table of the protostar properties.

Protostar Properties	
Temperature	4000 K
Radius	$0.9R_{\odot}$
Mass	$M \in [0.01, 0.2, 0.28]M_{\odot}$
Photons emitted	10^6 photons
Position in simulation grid	$(x, y, z) = (0, 0, 0)$

- **Alignment**

The alignment mechanism is given by Davis-Greenstein alignment theory [8] where paramagnetic relaxation creates a torque that causes the dust grain to rotate about its shorter axis. Refer to section 8 for more detail regarding grain alignment theories.

16 Simulation Results

As the core evolves, the magnetic field lines are extensively dragged inward to where they achieve a very pinched hourglass configuration. The vertical gradient in the B_φ component becomes very strong at the center and at the disk surface. The twisting in the field lines and creates a magnetic pressure gradient that drives outflows from the upper and lower surfaces of the outer disk and higher speed jets from the central region [47].

16.1 Synthetic Polarization Maps

The simulations are performed at all levels of the nested grid for $M = 0.01M_\odot$, $0.2M_\odot$ and $0.28M_\odot$, respectively. The simulation results are given in figures 44, 45 and 46. They showcase the level 12 synthetic polarization maps that resolve the polarization state of the small scale regions of the protostellar core at a series of declination angles. The declinations are given with respect to the $x - z$ plane, that is, the maps at $\theta = 0$ are assumed to be viewed face on by the observer. Each simulation has the outflow axis taken along the vertical axis (z -direction) and the disk midplane taken directly in the horizontal direction (x -direction). Both the polarization maps and figures 37 and 41 from section 14.4 indicate a toroidal structure in the magnetic field on small scales. This is indicative of rapid rotation and flux-freezing behavior on these scales. In the early stages of the core's evolution, the low ionization of the clouds reduced the coupling of the magnetic field to the neutral material. As the core develops, the compression on small scales increases. This in turn increases the temperature in these regions which increases the collision rate of particles. From this there is an increase in the fractional ionization which creates enough charged particles to cause flux freezing in the regions close to the protostar. This can be seen through a crude treatment of the Saha equation:

$$(16.1) \quad \frac{n_{i+1}n_e}{n_i} = \frac{2}{\lambda^3} \frac{g_{i+1}}{g_i} \exp \left[-\frac{(\epsilon_{i+1} - \epsilon_i)}{k_B T} \right]$$

where the left hand side represents the fractional ionization (fraction of an element in ionization-state $i + 1$ relative to state i), g_i is the degeneracy of states for i ions, ϵ_i is the required energy for an ion of level i , n_e is the density of electrons and λ is the thermal de Broglie wavelength given as:

$$\lambda = \sqrt{\frac{h^2}{2\pi m_e k_B T}}$$

If one raises the temperature T , it can be shown that the decaying exponential in equation 16.1 and the de Broglie wavelength will decrease. The decrease in these quantities will increase the fractional ionization and thus increase the number of charged particles, thereby causing an approach to flux freezing.

Mass: $0.01M_{\odot}$

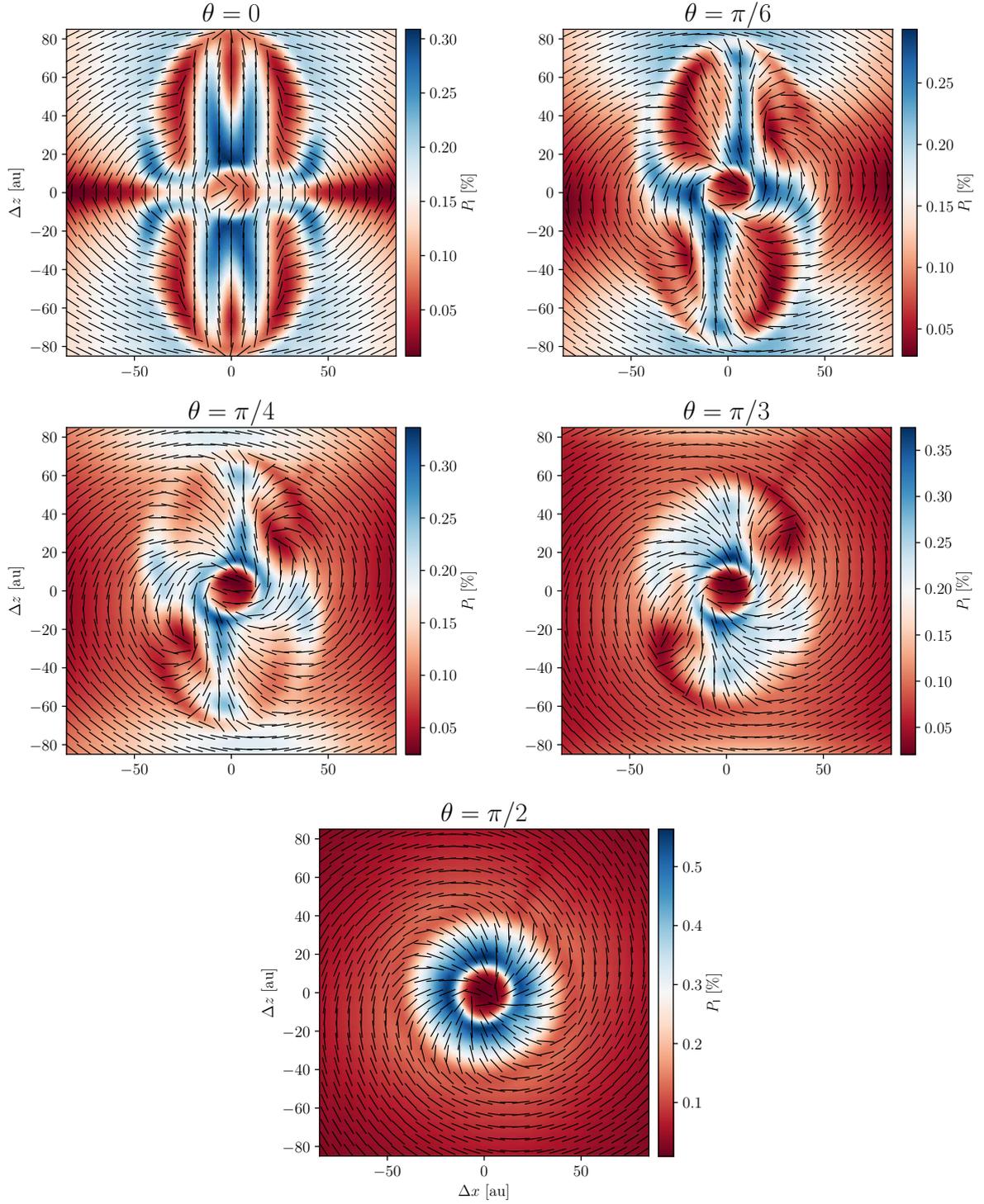


Figure 44: Synthetic polarization maps for $M = 0.01M_{\odot}$ demonstrated at a series of declination angles given in terms of the polarization efficiency P [%].

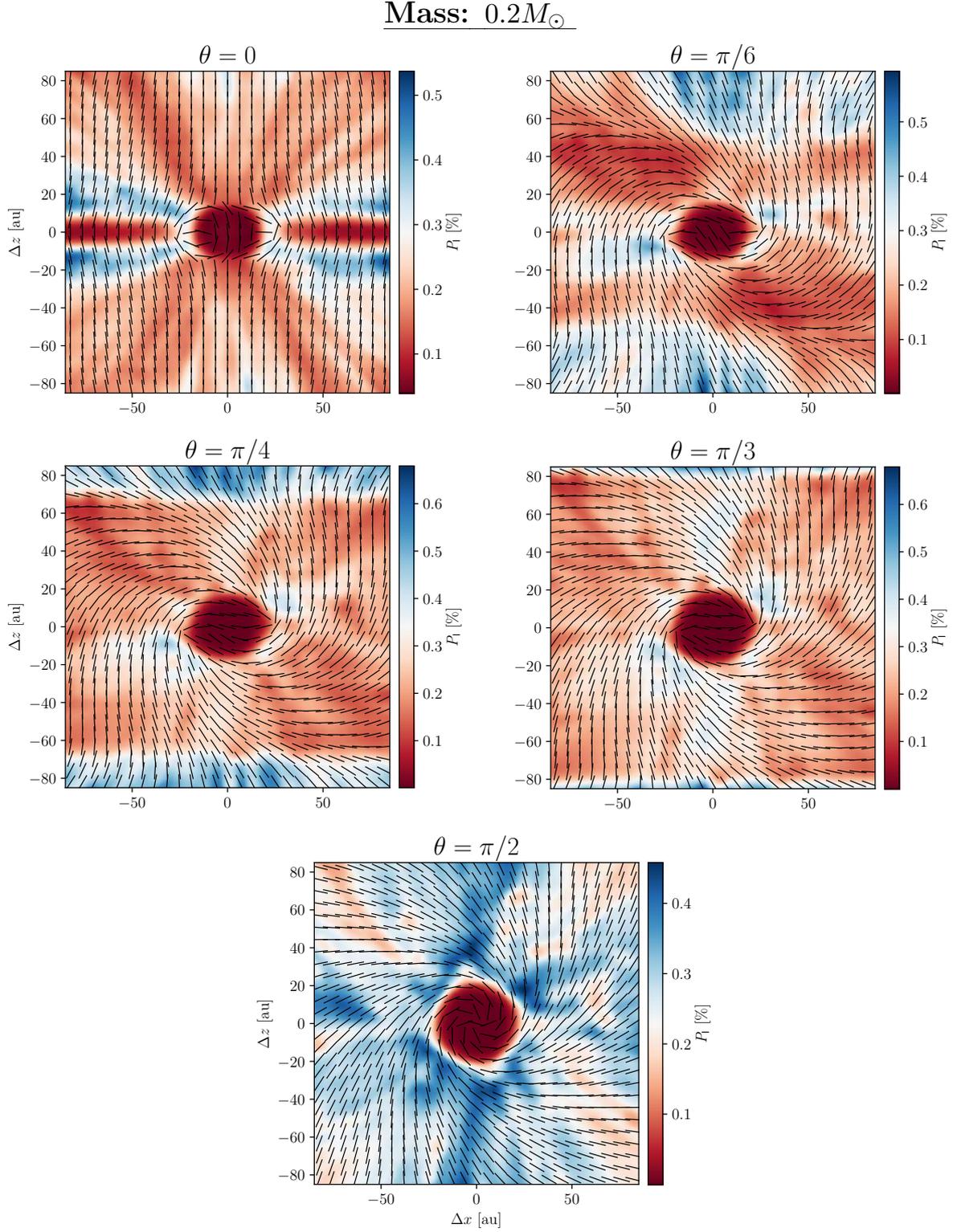


Figure 45: Synthetic polarization maps for $M = 0.2M_{\odot}$ demonstrated at a series of declination angles given in terms of the polarization efficiency P [%].

Mass: $0.28M_{\odot}$

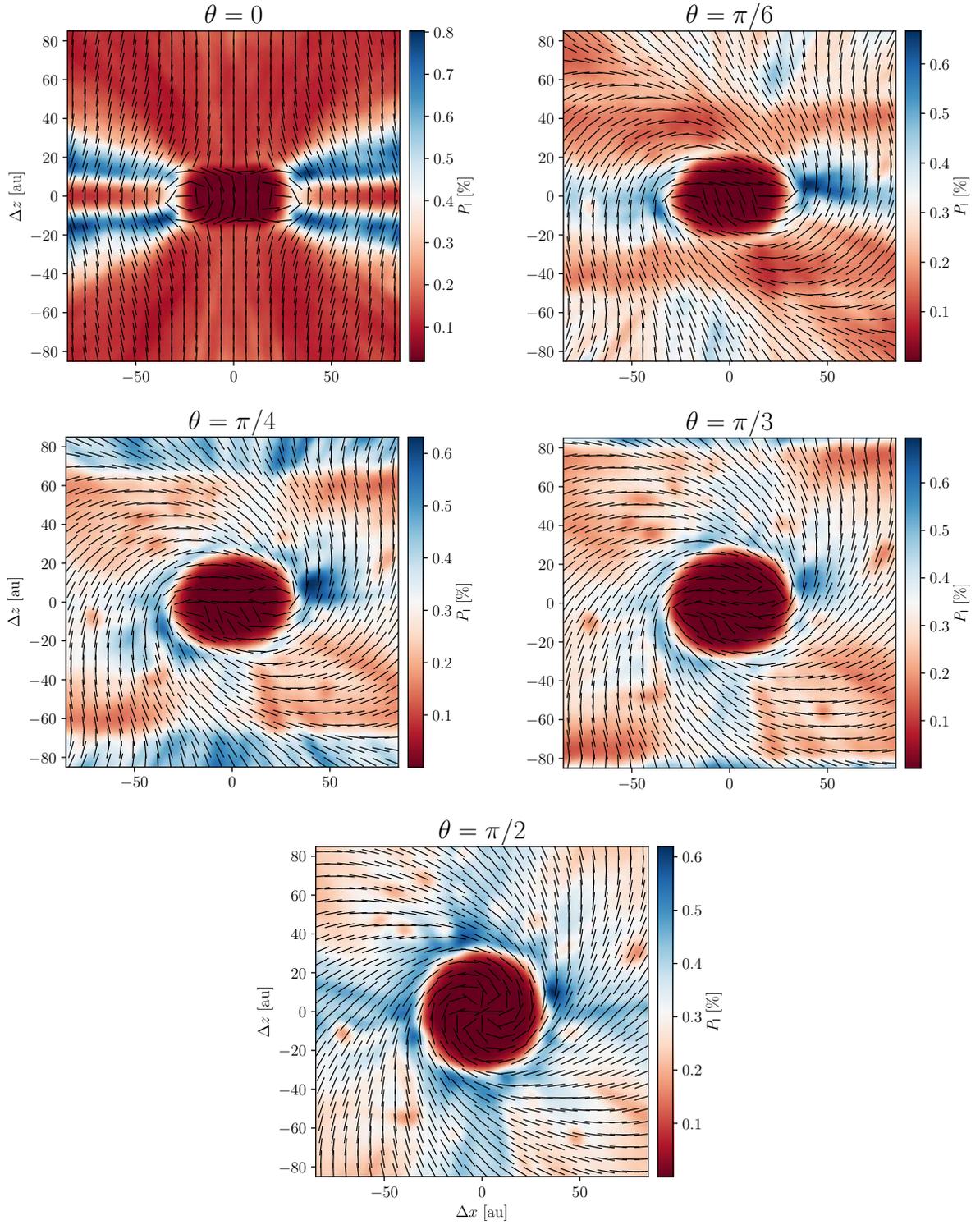


Figure 46: Synthetic polarization maps for $M = 0.28M_{\odot}$ demonstrated at a series of declination angles given in terms of the polarization efficiency P [%].

16.2 Modelling the Polarization of Orion Source I

The simulated synthetic polarization maps generated using POLARIS can be used in conjunction with observational data to model the polarization state and magnetic field configuration of the protostellar region known as Orion Source I [51]. Hirota et al. [51] probed the magnetic field in the bipolar outflow region of a high mass protostar in Orion Source I. Using the Atacama Large Millimeter/Submillimeter Array (ALMA), polarization measurements were taken of SiO rotational transitions. Hirota et al. [51] mapped the resulting linearly polarized light that was emitted in the ground state transitions of the $J = 1 - 0$ and $J = 2 - 1$ lines having frequencies of 43 GHz and 86 GHz respectively. These measurements had demonstrated varying polarization angles that Hirota et al. [51] speculate to be a result of Faraday rotation. Faraday rotation occurs when light and magnetic field interact through a medium. The magnetic field threading the partially ionized gas acts as a magnetoactive medium and when linearly polarized light travels through, its electric field component will exhibit a force on the free electrons. This will result in a circular motion of charged particles that will create a phase shift between the left and right handed circular polarization components and results in a rotation of the plane of linear polarization.[†] Figure 47 demonstrates the polarization vectors measured in Hirota et al. [51] which have been corrected for Faraday rotation. Using the DCF method, they estimated the magnetic field strength in the outflow lobe to be ≈ 30 mG with a configuration being either toroidal or poloidal depending on the relative orientation of polarization vectors with respect to the magnetic field. A toroidal structure would be consistent with magnetic field simulated from the MHD model (see figures 37 and 41 levels 11 - 13) and inferred from the POLARIS simulation. By matching the length scale of the POLARIS outputs directly with the ALMA measurements, the synthetic polarization maps can be used to map the magnetic field morphology and study the polarization vector alignment properties. However, there is uncertainty in whether polarization in the molecular spectral lines of SiO are parallel or perpendicular to the local magnetic field. Due to the high fractional polarization in their observations (suggestive of masers being anisotropically pumped), they cannot make inference on the alignment using the van Vleck angle relationship developed by Goldreich et al. [52]. That is, if masers[‡] are isotropically pumped, there exists a simple relationship with respect to the van Vleck angle[§] θ_v where if $\theta_v > 55^\circ$, the polarization is perpendicular to the magnetic field and if $\theta_v < 55^\circ$ the polarization is parallel to the magnetic field. POLARIS outputs the expected polarization direction in dust emission. These segments are oriented perpendicular to the mean magnetic field direction in the plane of the sky. If the SiO line emission has polarization segments parallel to the POLARIS segments, then the expected polarization mechanism for the SiO lines is producing polarization perpendicular to the local magnetic field. It was found that the level 10

[†]Linearly polarized light can be decomposed into the superposition of two circularly polarized components of opposite handedness and equal amplitude.

[‡]Stimulated spectral line emission in the microwave spectrum.

[§] θ_v is the angle made between the magnetic field and maser propagation direction.

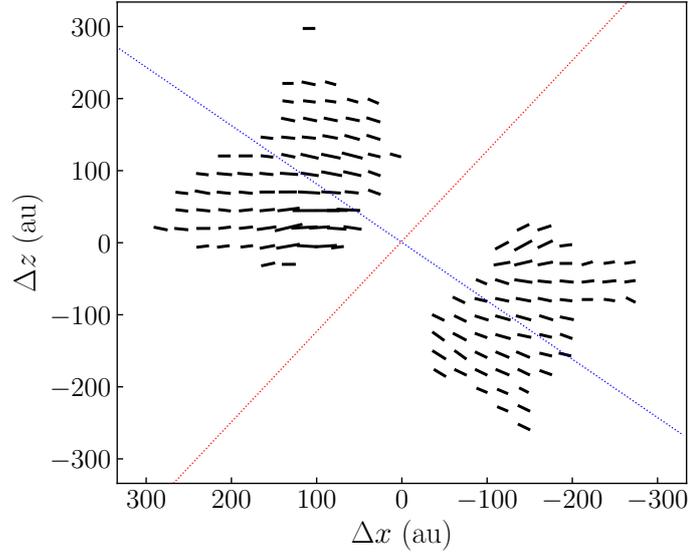


Figure 47: Polarization vector map corrected for Faraday rotation. Blue and red dashed lines indicate the position angle of the outflow axis (51 degrees) and disk midplane (141 degrees), respectively [51].

output from POLARIS best matches the length scales of the ALMA data and is given at a series of declination angles taken from $\theta \in [0^\circ, 10^\circ, 20^\circ, 30^\circ, 40^\circ, 45^\circ, 50^\circ, 60^\circ, 70^\circ, 80^\circ, 90^\circ]$. Figures 49, 50 and 51 demonstrate the superimposition of the polarization vectors from the ALMA observations with the synthetic polarization maps of POLARIS at $\theta \in [0, \pi/6, \pi/4, \pi/3, \pi/2]$. In each figure, the outflow axis of the POLARIS simulation has been adjusted to match the observed outflow axis in figure 47 (blue line). The alignment mechanism can be quantified by estimating the mean offset angle that each individual polarization vector makes with respect to the z -axis. That is, in both the POLARIS simulation and the ALMA observation, the quantity $0 \leq \phi \leq \pi/2$ is defined as the offset angle made between the polarization vector and the z -axis. Furthermore, the quantity $\Delta\hat{\phi}$ is the absolute residual between the mean of the offset angles from the ALMA observation and the POLARIS simulation. That is,

$$(16.2) \quad \Delta\hat{\phi}^{(l)} = \left| \hat{\phi}_A - \hat{\phi}_P^{(l)} \right|$$

having $\hat{\phi}$ be the mean offset angle given by:

$$(16.3) \quad \hat{\phi} = \frac{1}{N} \sum_i^N \phi_i$$

where N is the total number of polarization vectors. In equation 16.2, the superscript (l) is taken to refer to the nested grid level in the POLARIS simulation and the subscripts A and P indicate the mean offset angle taken for the ALMA observation and the POLARIS output, respectively. By default POLARIS will output polarization approximately perpendicular to the local magnetic field. Thus in order to truly assess the alignment mechanism in the ALMA observation, a second

calculation will be performed where the polarization vectors from POLARIS will be rotated by 90 degrees in order to infer if the ALMA polarization is parallel to the local magnetic field. Table 16.1 gives the calculations for $\Delta\hat{\phi}^{(10)}$ for POLARIS polarization perpendicular $\Delta\hat{\phi}_{\perp}^{(10)}$ and parallel $\Delta\hat{\phi}_{\parallel}^{(10)}$ to the local magnetic field. The inference is made by taking the the smallest residual which is given for a declination angle $\theta = 90^\circ$ with polarization vectors parallel to the local magnetic field. Although this model presents the smallest residual of $\Delta\hat{\phi}_{\parallel} = 11.67^\circ$, the observation suggests that the disk is approximately edge-on where as a declination of $\theta = 90^\circ$ is indicative of a head-on view. Thus, the best representative model is taken to be that having a declination of $\theta = 40^\circ$ with a residual $\Delta\hat{\phi}_{\parallel} = 12.00^\circ$. This residual is not far off from 11.67° and the critical result is that both models imply that the SiO line emission polarization vectors are parallel to the local magnetic field. Figure 48 gives the synthetic maps output from POLARIS for $\theta = 40^\circ$. Additionally, the right panel of figure 48 also demonstrates the inferred magnetic field morphology.

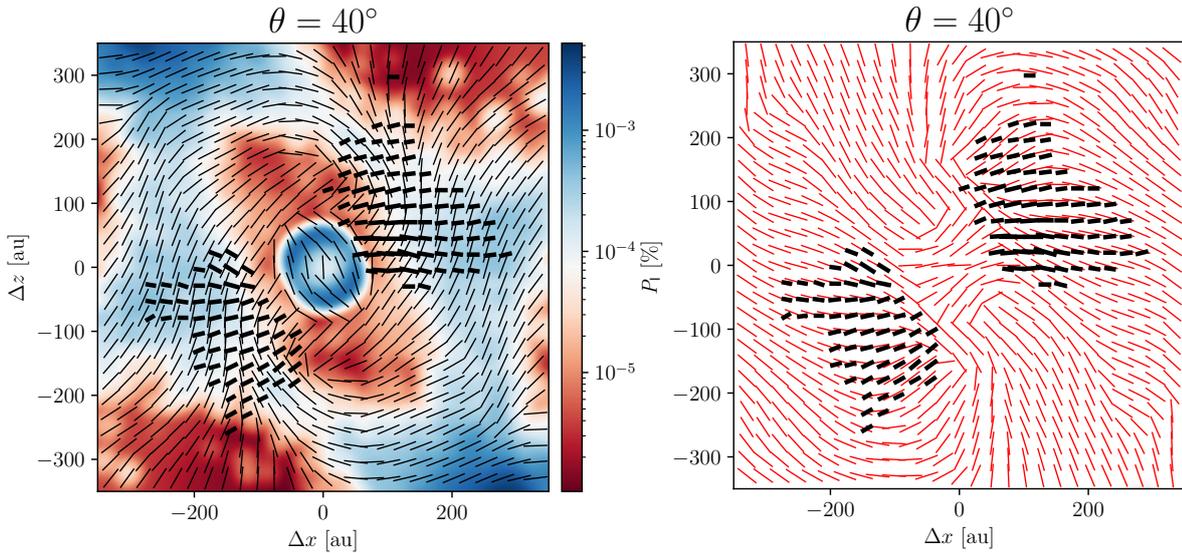


Figure 48: POLARIS simulation output for the best representative model adopting $M = 0.28M_{\odot}$ at declination of $\theta = 40^\circ$. **Left:** Simulated synthetic polarization map from POLARIS overlaid with the polarization segments from the ALMA data (bold vectors). **Right:** Inferred magnetic field morphology from the POLARIS output assuming parallel alignment (red vectors) overlaid with the ALMA polarization segments. The magnetic field is inferred by rotating the polarization vectors in the left panel by 90 degrees.

Table 16.1: Table representative of the respective mean offset angles for each mass instance. The quantity $\Delta\phi_{\perp}^{(10)}$ is the absolute residual angle between the POLARIS simulation and the ALMA data assuming perpendicular alignment. The quantity $\Delta\phi_{\parallel}^{(10)}$ is the absolute residual angle between the POLARIS simulation and the ALMA data assuming parallel alignment.

θ (degrees)	$M = 0.01M_{\odot}$	$M = 0.2M_{\odot}$	$M = 0.28M_{\odot}$
	$\Delta\phi_{\perp}^{(10)}$ (degrees)		
$\theta = 0^{\circ}$	44.38 $^{\circ}$	32.90 $^{\circ}$	32.51 $^{\circ}$
$\theta = 10^{\circ}$	37.49 $^{\circ}$	38.09 $^{\circ}$	43.08 $^{\circ}$
$\theta = 20^{\circ}$	49.62 $^{\circ}$	44.68 $^{\circ}$	48.53 $^{\circ}$
$\theta = 30^{\circ}$	48.31 $^{\circ}$	36.11 $^{\circ}$	52.01 $^{\circ}$
$\theta = 40^{\circ}$	41.96 $^{\circ}$	47.64 $^{\circ}$	54.66 $^{\circ}$
$\theta = 45^{\circ}$	52.52 $^{\circ}$	34.25 $^{\circ}$	55.00 $^{\circ}$
$\theta = 50^{\circ}$	44.22 $^{\circ}$	44.23 $^{\circ}$	54.62 $^{\circ}$
$\theta = 60^{\circ}$	55.44 $^{\circ}$	38.75 $^{\circ}$	56.27 $^{\circ}$
$\theta = 70^{\circ}$	49.62 $^{\circ}$	48.52 $^{\circ}$	56.77 $^{\circ}$
$\theta = 80^{\circ}$	52.28 $^{\circ}$	51.65 $^{\circ}$	52.71 $^{\circ}$
$\theta = 90^{\circ}$	54.38 $^{\circ}$	48.69 $^{\circ}$	46.35 $^{\circ}$
	$\Delta\phi_{\parallel}^{(10)}$ (degrees)		
$\theta = 0^{\circ}$	26.80 $^{\circ}$	29.05 $^{\circ}$	28.74 $^{\circ}$
$\theta = 10^{\circ}$	25.00 $^{\circ}$	23.30 $^{\circ}$	20.10 $^{\circ}$
$\theta = 20^{\circ}$	14.86 $^{\circ}$	19.23 $^{\circ}$	17.31 $^{\circ}$
$\theta = 30^{\circ}$	23.65 $^{\circ}$	26.15 $^{\circ}$	14.30 $^{\circ}$
$\theta = 40^{\circ}$	20.80 $^{\circ}$	18.62 $^{\circ}$	12.00 $^{\circ}$
$\theta = 45^{\circ}$	19.80 $^{\circ}$	28.73 $^{\circ}$	12.72 $^{\circ}$
$\theta = 50^{\circ}$	18.80 $^{\circ}$	22.98 $^{\circ}$	13.21 $^{\circ}$
$\theta = 60^{\circ}$	14.55 $^{\circ}$	23.37 $^{\circ}$	13.02 $^{\circ}$
$\theta = 70^{\circ}$	14.86 $^{\circ}$	20.31 $^{\circ}$	12.33 $^{\circ}$
$\theta = 80^{\circ}$	13.22 $^{\circ}$	16.44 $^{\circ}$	14.84 $^{\circ}$
$\theta = 90^{\circ}$	11.67 $^{\circ}$	16.46 $^{\circ}$	17.75 $^{\circ}$

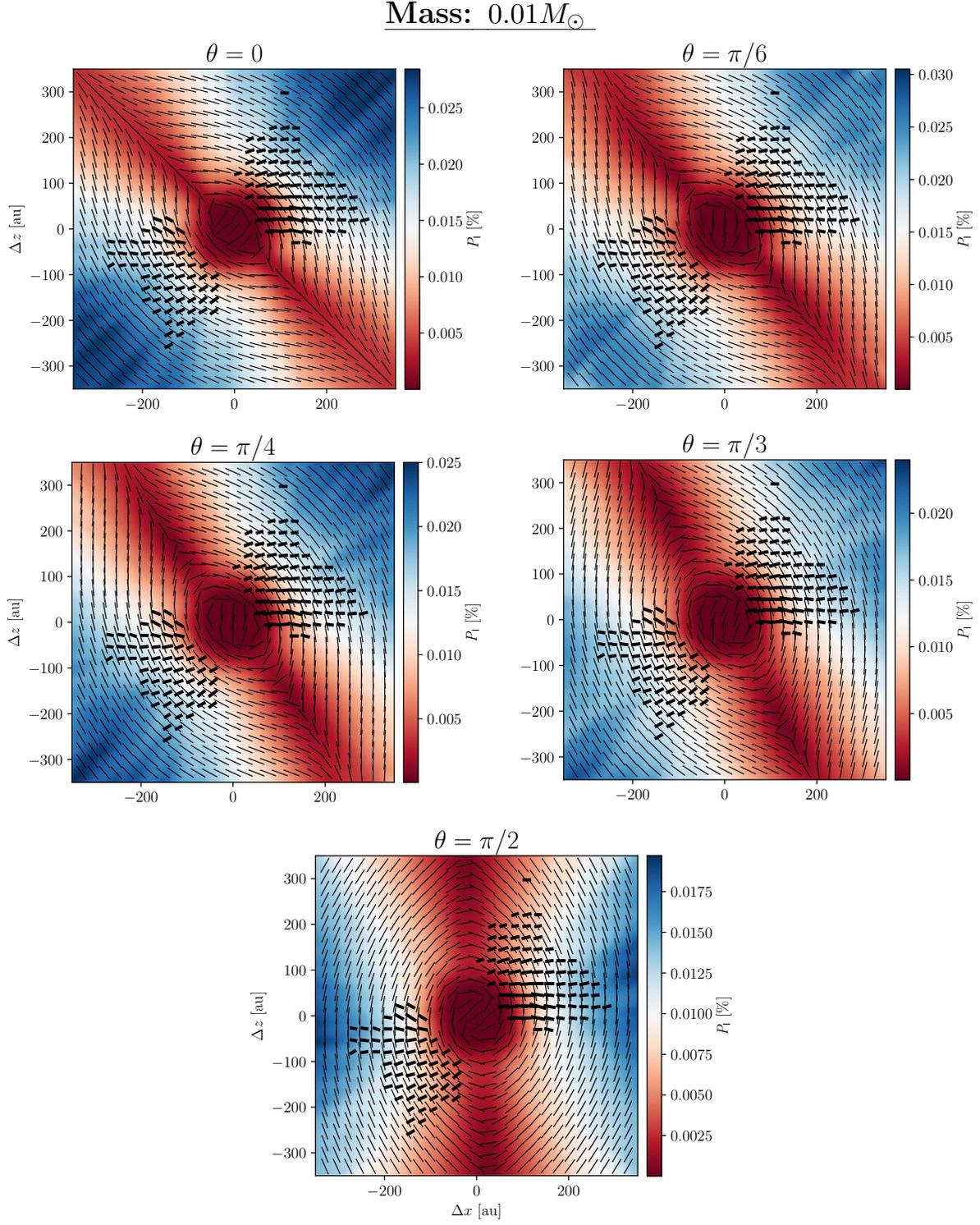


Figure 49: The ALMA data overlaid with the simulated synthetic polarization maps for $M = 0.01M_{\odot}$.

Mass: $0.2M_{\odot}$

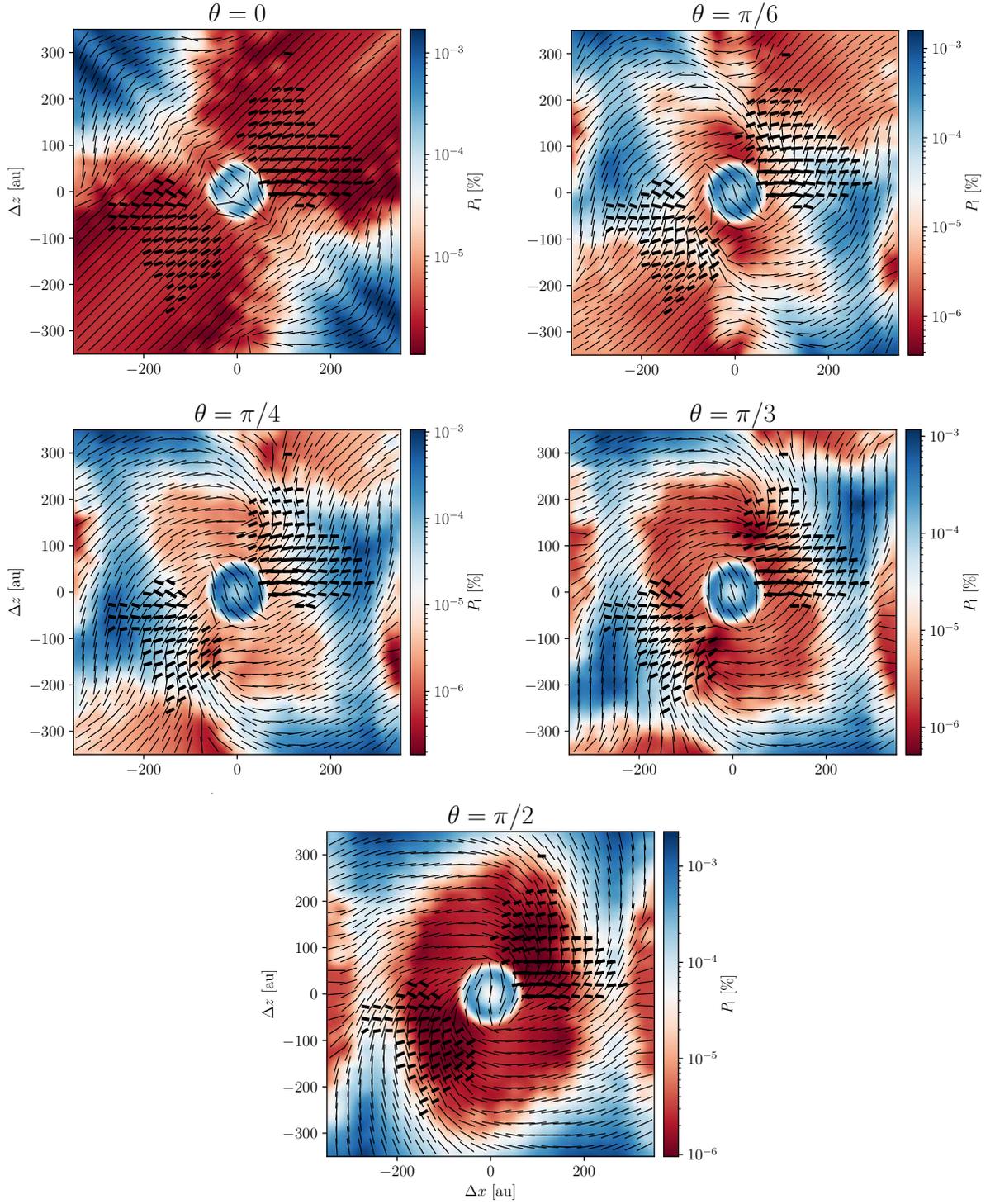


Figure 50: The ALMA data overlaid with the simulated synthetic polarization maps for $M = 0.2M_{\odot}$.

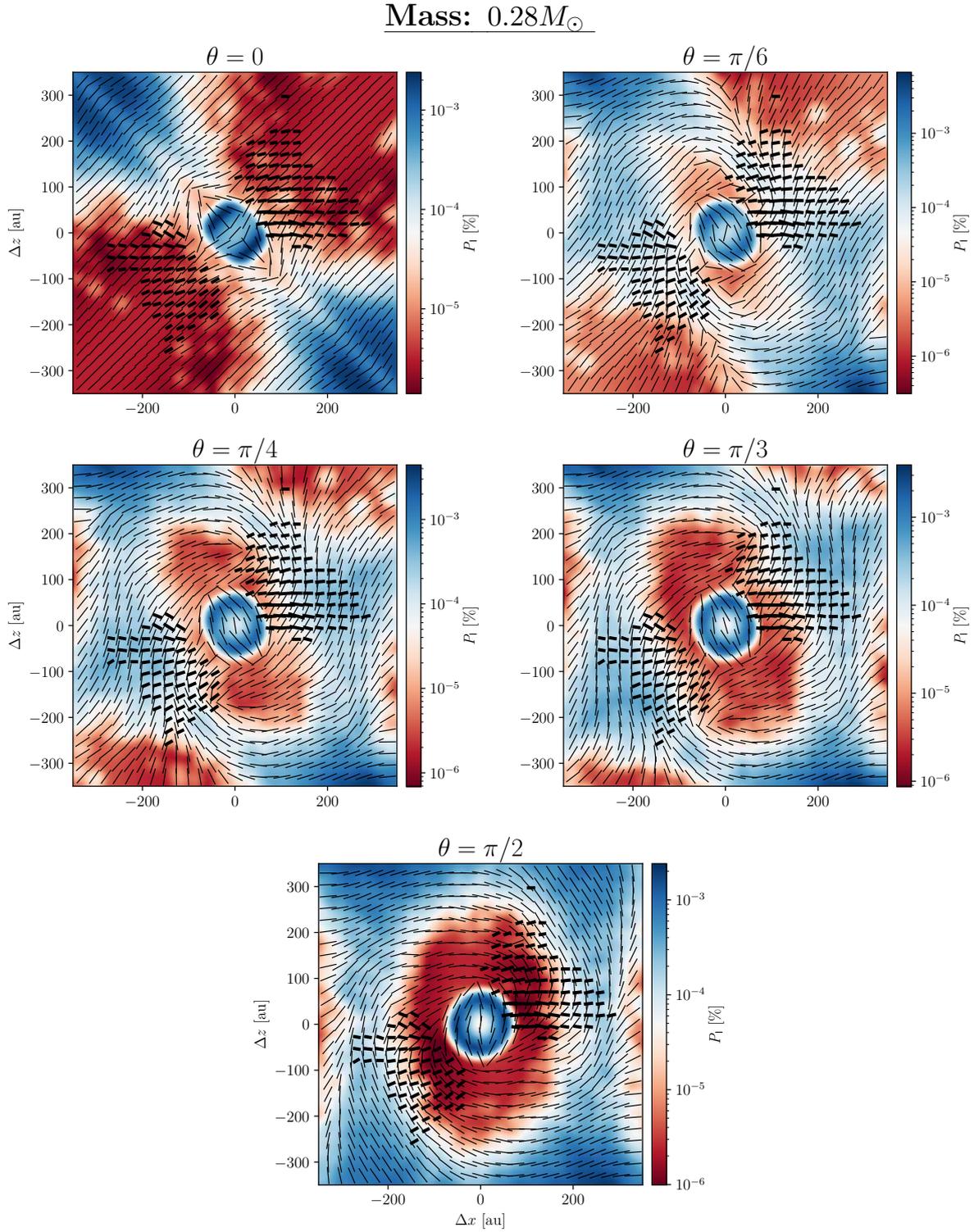


Figure 51: The ALMA data overlaid with the simulated synthetic polarization maps for $M = 0.28M_{\odot}$.

17 Conclusion

By utilizing non-ideal MHD simulations to model the properties of protostellar cores, simulated synthetic polarization maps were used to model the polarization state of protostellar cores at a series of length scales and declination angles. By implementing the MHD simulations on a nested grid, the POLARIS outputs can demonstrate high resolution polarization maps for length scales as small as several astronomical units. Using this in conjunction with a series of declination angles, the synthetic maps can be used to make inference on the polarization state of Orion Source I. By superimposing the level 10 POLARIS outputs with the ALMA observational data and deploying a calculation on the residual offset angle made with respect to the z -axis, it was found that the data is best modelled at a declination angle of roughly 40° . A good fit is obtained if the observed molecular spectral line polarization is parallel to the local magnetic field. The model helps to settle the ambiguity about the alignment of the polarization in relation to the magnetic field.

EPISODIC ACCRETION & MASS ACCUMULATION

Collaborator: Lyle Muller, University of Western Ontario

Key Words: accretion: young stellar objects – time series: neural networks

18 Introduction

As a protostellar core accrues more gas from its surroundings, the initial collapse begins with a given amount of angular momentum. Since the angular momentum is conserved, the rotation will increase causing the infalling gas to eventually flatten and form an accretion disk. The centrifugal acceleration from the rotation will resist the radial gravitational force from the central object, but collapse can still occur along the vertical direction. Continuous accumulation of mass can however lead to an unstable disk that results in a series of bursts during the accretion phase. Modelling such dynamics can be done by solving the hydrodynamical equations for mass, momentum, and energy transport. The output is typically a highly chaotic time series consisting of several bursts (or spikes) representative of the episodic behavior experienced throughout the instability. Modelling chaotic systems possess the feature of unpredictability. That is, the dynamics are severely sensitive to the initial conditions. If two initial conditions differ by an amount δx , after a time t they will differ by $\delta x e^{\Lambda t}$ experiencing exponential separation. Most chaotic systems cannot be directly modelled by dynamical systems and are only available through time series. There are many statistical models available for time series such as the Autoregressive Integrated Moving Average (ARIMA), Simple Exponential Smoothing (SES) and Vector Autoregression (VAR) models to name a few. Additionally, artificial forms of recurrent neural networks such as LSTM and ESN models have demonstrated effective modelling of time series. ESN models have particularly demonstrated effective prediction of highly chaotic systems for longer periods of time than other deep learning algorithms [53]. ESN models do however present some statistical challenges. Several alternatives have been demonstrated to mitigate such challenges utilizing ensemble methods and parallel series architectures [54, 55]. These methodologies can be extended to forming a constraint optimization problem for a set of weights to give an optimally weighted alternative (OWA-ESN) to the parallel series method. This approach is introduced as an effective, model-free technique for high chaotic time series. The model robustness can be tested on the hydrodynamic simulations to demonstrate the ability of extracting the underlying temporal dynamics and providing effective physical interpretation of the model outputs for a series of initial conditions.

19 Hydrodynamics & Stellar Evolution

Simulating stellar evolution begins with the collapse of a gravitationally contracting core and continues into the phase of star formation. This involves solving the hydrodynamical equations:

$$\begin{aligned} \frac{\partial \Sigma}{\partial t} &= -\nabla_{\mathbf{p}} \cdot (\Sigma v_{\mathbf{p}}) \\ \frac{\partial}{\partial t}(\Sigma v_{\mathbf{p}}) + [\nabla \cdot (\Sigma v_{\mathbf{p}} \otimes v_{\mathbf{p}})]_{\mathbf{p}} &= -\nabla_{\mathbf{p}} P + \Sigma g_{\mathbf{p}} + (\nabla \cdot \Pi)_{\mathbf{p}} \\ \frac{\partial e}{\partial t} + \nabla_{\mathbf{p}} \cdot (e v_{\mathbf{p}}) &= -P(\nabla_{\mathbf{p}} \cdot v_{\mathbf{p}}) - \Omega + \Gamma + (\nabla \cdot v)_{\text{pp}}, : \Pi_{\text{pp}}, \end{aligned}$$

which correspond to mass, momentum, and energy transport respectively. The equations define Σ to be the surface mass density, e is the internal energy per surface area, P is the vertically integrated pressure, $v_{\mathbf{p}} = v_r \hat{\mathbf{r}} + v_{\phi} \hat{\boldsymbol{\phi}}$ is the velocity over the plane of the disk and $\nabla_{\mathbf{p}} = \hat{\mathbf{r}} \partial / \partial r + \hat{\boldsymbol{\phi}} r^{-1} \partial / \partial \phi$ is the planar gradient operator that takes the gradient along the plane of the disk. Vorobyov et al. [56] define the gravitational acceleration in the disk plane $g_{\mathbf{p}}$ to take into account the self gravity of the disk and the gravity of the formed protostar [57]. Lastly, Π is the viscous stress tensor, the radiative cooling Ω and the heating function Γ are taken to be:

$$\Omega = F_c \sigma T_{\text{mp}}^4 \frac{\tau}{1 + \tau^2}, \quad \Gamma = F_c \sigma T_{\text{irr}}^4 \frac{\tau}{1 + \tau^2}$$

where σ is the Stefan-Boltzmann constant, T_{mp} is the midplane temperature of gas, T_{irr} is the irradiation temperature and τ is the optical depth. The large episodes of accretion are due to clumps that form within the disk due to gravitational instability and migrate to the center of the core. Further details regarding the simulations are found in Vorobyov et al. 2017 [56]. Accretion data will be used from 7 of the 35 models [56] where initial conditions are found in Vorobyov et al. 2017 [56].

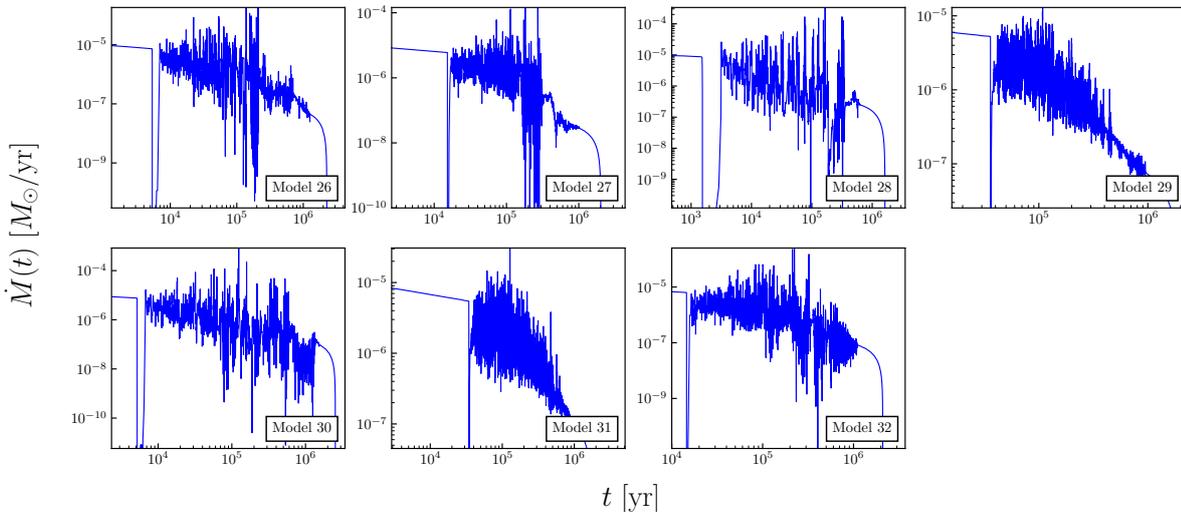


Figure 52: Hydrodynamic simulation outputs for 7 of the 35 models in Vorobyov et al. [56] (labeled 26-32).

19.1 Initial State

In this section, a broad description regarding the initial state of the models are discussed for completeness, but the reader should refer to Vorobyov et al. [56] for a more detailed discussion regarding the initial conditions used in the simulations. Each of the models that are used to simulate the accretion data are characterized by the initial radial profiles for the gas surface density Σ and angular velocity ω taking the form

$$\Sigma = \frac{r_0 \Sigma_0}{\sqrt{r^2 + r_0^2}}, \quad \omega = 2\omega_0 \left(\frac{r_0}{r}\right)^2 \left[\sqrt{1 + \left(\frac{r}{r_0}\right)^2} - 1 \right]$$

for Σ_0 and ω_0 being the surface mass density and angular velocity at the center of the core and $r_0 = \sqrt{Ac_s^2/\pi G\Sigma_0}$ being the central plateau radius adopting $A = 1.2$ [56]. The model is two-dimensional, assumes no magnetic field and takes the initial gas temperature in collapsing cores to be 10 K. In order to generate gravitationally unstable truncated cores, each model is characterized by the ratio $r_{\text{out}}/r_0 = 6$ where r_{out} is the core's outer radius. Once r_{out} has been selected, r_0 can be found via the ratio set between the two quantities and the central surface density can then be found using $r_0 = \sqrt{Ac_s^2/\pi G\Sigma_0}$. Finally, the cloud core mass M_{cl} can be found using the initial radial profile for the gas surface density Σ . The quantity ω_0 is selected such that the ratio of rotational to gravitational energy β falls within $\sim 10^{-4}$ and 0.07 [58]. The initial parameters are summarized in table 19.1.

Table 19.1: Initial parameters for models 26-32. The quantity $M_{*,\text{fin}}$ is the final stellar mass.

Model	$M_{\text{core}}[M_{\odot}]$	$\beta[\%]$	$r_0[\text{au}]$	$M_{*,\text{fin}}[M_{\odot}]$
Model 26	1.245	1.27	2777	0.753
Model 27	1.076	0.56	2400	0.801
Model 28	1.767	2.25	3943	0.807
Model 29	0.999	0.28	2229	0.818
Model 30	1.537	1.27	3429	0.887
Model 31	1.306	0.28	2915	1.031
Model 32	1.383	0.56	3086	1.070

20 Modelling Chaotic Time Series

The focal point of this section involves demonstrating the predictive capabilities of neural networks with applications to the episodic accretion simulations introduced. This section begins with a discussion on chaotic time series and the data pre-processing methodology and builds up to introducing the development and deployment of a specialized form recurrent neural networks implementing the reservoir computing framework and constraint optimization. A reader who is interested in the application of the methodology can go directly to section 20.7.

20.1 Chaos & Maximum Lyapunov Exponent

The dynamics of the accretion rate in figure 52 demonstrate what is believed to be very chaotic behavior. This chaos can be quantified in a constant Λ known as the Lyapunov exponent. The Lyapunov exponent effectively characterizes the rate of separation between close trajectories in phase space. Typically, the maximum Lyapunov exponent is taken to characterize a time series and the data can be interpreted as:

$$\begin{aligned}\Lambda_{\max} > 0, & \quad \text{Chaotic Dynamics} \\ \Lambda_{\max} = 0, & \quad \text{Regular Dynamics} \\ \Lambda_{\max} < 0, & \quad \text{Fixed-Point Dynamics}\end{aligned}$$

where Λ_{\max} is taken to be the maximum Lyapunov exponent. In phase space, two signals that differ by an amount $\delta\mathbf{X}_0$ will diverge at a rate given by:

$$(20.1) \quad |\delta\mathbf{X}(t)| \sim e^{\Lambda t} |\delta\mathbf{X}_0|$$

In order to estimate the Lyapunov exponents in the mass accretion data, the Eckmann algorithm is used and the details are found in Eckmann & Kamphorst 1986 [59] (the reader is referred to appendix A1.5 for a brief walkthrough on the algorithm). The maximum Lyapunov exponents are calculated for each model and given in table 20.1. Given that the mass accretion rate calms to a constant value after the chaotic period, the analysis will only be performed on the episodic portion of each respective model.

Table 20.1: Maximum Lyapunov exponent for each respective model.

Model	Λ_{\max}
Model 26	0.20821945
Model 27	0.2014386
Model 28	0.20973413
Model 29	0.18445392
Model 30	0.16244096
Model 31	0.19877397
Model 32	0.1922906

20.2 Neural Networks

Neural networks have been demonstrated to accurately approximate any continuous function and are often referred to as universal approximators. This ability has given neural networks tremendous capabilities to extract and estimate the underlying dynamical process occurring for a time series. In its most vanilla form, neural networks feed input data from an input layer into one or many hidden layers before a prediction is made at the output layer. Each layer is made up of a series of neurons that contain data and their respective transformations from one layer to the next described by a set of weights. The single layer perceptron is often referred to as the fundamental unit of a neural network and its architecture schematically demonstrated in figure 53. The perceptron is a binary linear classifier where the weights are recursively updated

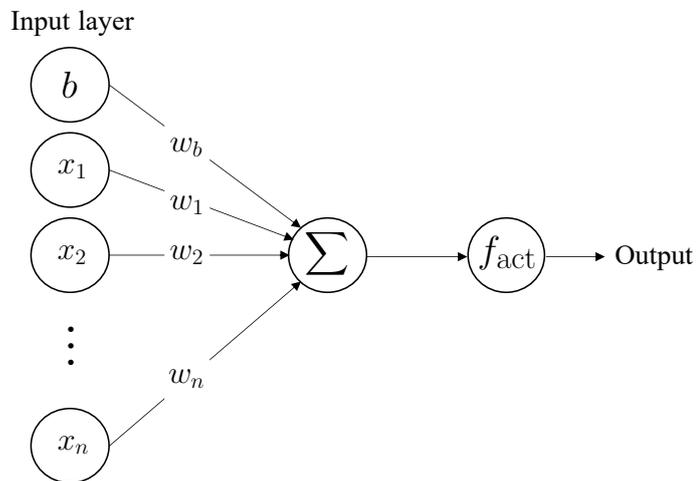


Figure 53: Schematic figure of a one layer perceptron. The input data $\mathbf{x} = [b, x_1, x_2, \dots, x_n]$ along with a bias b are given the weights $\mathbf{W} = [w_b, w_1, w_2, \dots, w_n]$, are summed and fed through an activation function f_{act} to return an output.

after each output until the iteration error $\varepsilon^{(i)}$ is below a user defined threshold. The algorithm for the one layer perceptron having N examples in a set $\mathbf{D}_{N \times n}$ used for training is given as:

1. Initialize the weights and threshold. Typically, weights are initialized to zero or a small random number.
2. For every j in $\mathbf{D}_{N \times n}$ containing a target output d_j :

(i) Compute the vector product of input layer data and the weight matrix:

$$g_j^{(i)} = (\mathbf{W}^{(i)})^T \cdot \mathbf{x}_j^{(i)}$$

(ii) Feed $g^{(i)}$ into an activation function for an output $y^{(i)}$:

$$y_j^{(i)} = f_{\text{act}}(g_j^{(i)})$$

(iii) Update the weights for a given learning rate γ :

$$\mathbf{W}^{(i+1)} = \mathbf{W}^{(i)} + \gamma \cdot (d_j - y_j^{(i)})\mathbf{x}_j^{(i)}$$

3. Define an error metric and compute the error as:

$$\varepsilon^{(i)} = \frac{1}{N} \sum_{j=1}^N |d_j - y_j^{(i)}|$$

4. If the error is above the user defined threshold, return to step 2 and repeat until convergence is achieved.

Given the limitations of a linear model like the single layer perceptron, it is natural to elevate to non-linear models like the multi-layer perceptron or more generally, the multilayered neural network. Given that this analysis deals primarily with time series, assume that the input data now consists of a time series $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$. The multilayered network architecture is given in figure 54 and demonstrates a feed forward network where the inputs are fed through the network in one direction. The concept is similar to the single layer perceptron

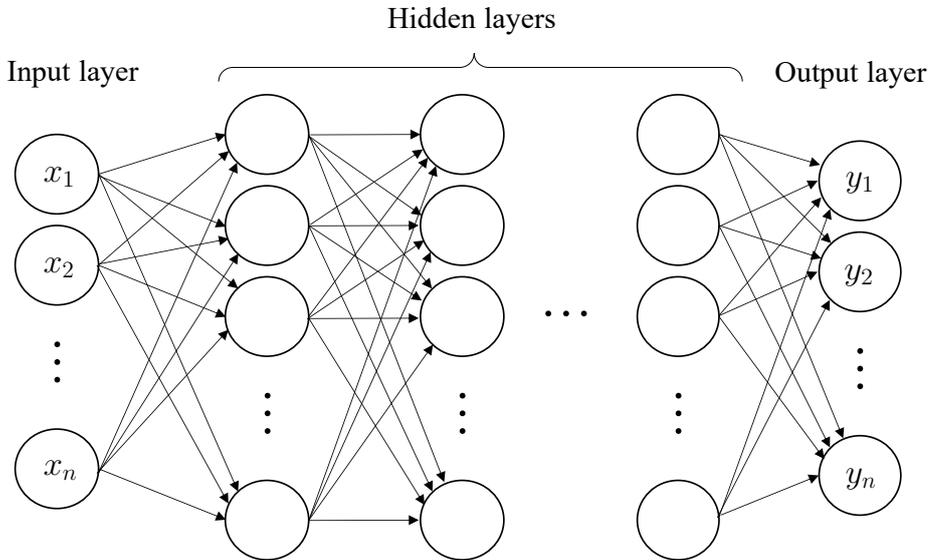


Figure 54: Schematic figure of a multilayered neural network.

where the connectivity of each layer is assigned a matrix $\mathbf{W}^{(l)}$ that represents the weights connecting a given network layer l . The hidden layers apply non-linear transformations to inputs through activation functions and propagate the data forward through the network. The application to time series will have each neuron output a predictor y_i that estimates the value x_{i-1} in the time series. That is, for a time series input $\mathbf{x}(t)$, the network outputs the predictors $[y_1(t), y_2(t), \dots, y_n(t)]^T \equiv [\hat{x}_2(t), \hat{x}_3(t) \dots, \hat{x}_{n+1}(t)]^T = \hat{\mathbf{x}}(t)$, where the value $\hat{x}_i(t)$

is the estimate to the input value $x_i(t)$. Similar to the single layer perceptron, data is fed forward through the network through a forward propagation methodology and an error metric is calculated to quantify the network's performance. The forward propagation algorithm is typically given as:

$$\mathbf{h}^{(l+1)} = f_{\text{act}}(\mathbf{z}^{(l+1)}) = f_{\text{act}}((\mathbf{W}^{(l)})^T \cdot \mathbf{h}^{(l)})$$

for $\mathbf{h}^{(l)}$ being the output at layer l where $\mathbf{h}^{(1)} = \mathbf{x}(t)$ and $\mathbf{h}^{(L)} = \hat{\mathbf{x}}(t)$ for a network having L layers. At the output layer, the mean-squared error (or some other loss) is given as:

$$(20.2) \quad \text{MSE} = \sum_i \frac{(\hat{x}_i - x_i)^2}{N}$$

where N is the length of the input vector. If the error is above a user defined threshold, the weights are recursively updated according to a backpropagation algorithm. This typically involves changing the weights to minimize a cost function, C given in terms of some error metric. In order to minimize, the derivative of the cost function is taken with respect to the input:

$$(20.3) \quad \frac{dC}{d\mathbf{a}^{(L)}} \cdot \frac{d\mathbf{a}^{(L)}}{d\mathbf{z}^{(L)}} \cdot \frac{d\mathbf{z}^{(L)}}{d\mathbf{a}^{(L-1)}} \cdot \frac{d\mathbf{a}^{(L-1)}}{d\mathbf{z}^{(L-1)}} \cdot \frac{d\mathbf{z}^{(L-1)}}{d\mathbf{a}^{(L-2)}} \cdots \frac{d\mathbf{a}^{(1)}}{d\mathbf{z}^{(1)}} \cdot \frac{\partial \mathbf{z}^{(1)}}{\partial \mathbf{x}(t)}$$

where $\mathbf{a}^{(l)} = f_{\text{act}}(\mathbf{z}^{(l)})$. Therefore, this is written as:

$$\nabla_x C = (\mathbf{W}^{(1)})^T (\mathbf{a}^{(1)})' \cdots (\mathbf{W}^{(L-1)})^T (\mathbf{a}^{(L-1)})' \cdot (\mathbf{W}^{(L)})^T (\mathbf{a}^{(L)})' \cdot \frac{dC}{d\mathbf{a}^{(L)}}$$

Defining the quantity $\boldsymbol{\sigma}^{(l)}$ as:

$$\boldsymbol{\sigma}^{(l)} = (\mathbf{W}^{(l)})^T (\mathbf{a}^{(l)})' \cdots (\mathbf{W}^{(L-1)})^T (\mathbf{a}^{(L-1)})' \cdot (\mathbf{W}^{(L)})^T (\mathbf{a}^{(L)})' \cdot \frac{dC}{d\mathbf{a}^{(L)}}$$

and the derivative of the cost function with respect to the weights $\mathbf{W}^{(l)}$ is:

$$\frac{\partial C}{\partial \mathbf{W}^{(l)}} = \boldsymbol{\sigma}^{(l)} \cdot (\mathbf{a}^{(l-1)})^T$$

and so the weights are updated as:

$$\mathbf{W}_{j+1}^{(l)} = \mathbf{W}_j^{(l)} - \gamma \frac{\partial C}{\partial \mathbf{W}_j^{(l)}}$$

for a learning rate γ . To further generalize the network architecture, one can introduce a series of feedback loops within the network to allow information to move in both directions. These types of networks are called recurrent neural networks (RNN) and have the property of retaining 'memory' of previous input states. RNN's are well suited for time series because the internal state of the network is maintained from one time step to another. The basic architecture is given in figure 55 and demonstrates the feedback sequence for a basic 3-layered RNN. RNN's will serve as the foundation for the echo state neural network (ESN) used in this analysis.

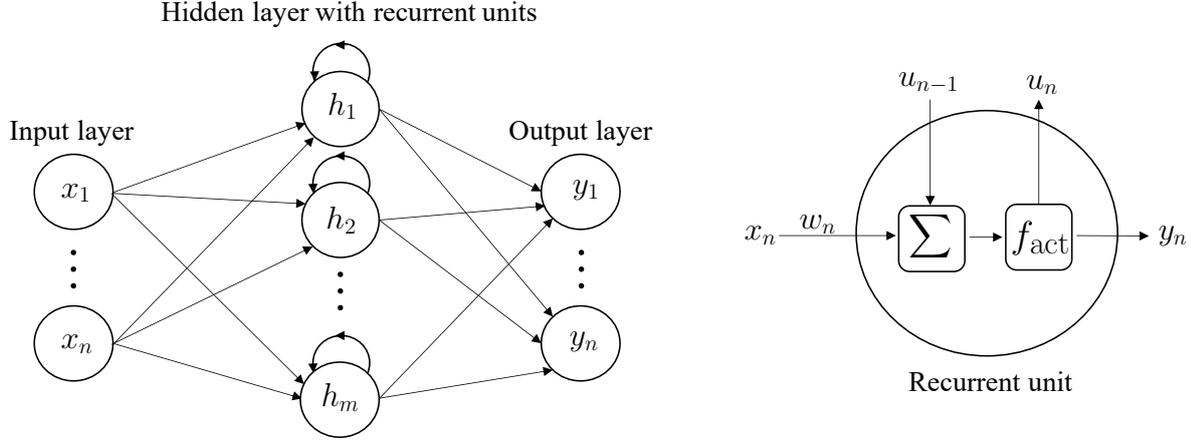


Figure 55: 3-layer RNN where the hidden layer neurons are recurrent and feed context outputs u_i back into the neuron in the form of 'memory'. The figure on the right demonstrates the recurrent sequence that each unit h_i outputs.

20.3 Reservoir Computing & Echo State Neural Networks

The reservoir computing framework is critical to the functionality benefits of echo state neural networks. The framework maps an input signal to a higher dimensional computational space. The reservoir consists of sparse, randomly connected neurons with a user defined connectivity having fixed, non-linear dynamics. One of the key benefits comes from computational efficiency. Unlike a traditional neural network that trains through a series of forward and backward propagation through the network, the reservoir computing framework only requires output weights to be trained. An echo state neural network is made up of an input layer, a reservoir and an output layer having weight matrices \mathbf{W}_{input} , \mathbf{W}_x and \mathbf{W}_{output} respectively. The traditional network architecture is demonstrated in figure 56. Like any neural network, in order to make predictions the ESN must be trained and validated. Assume that m elements of the time series $\mathbf{x}(t)$ are used for training and given by $\mathbf{x}_{train}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ for $m < n$ where n is the total number of data points in the time series $\mathbf{x}(t)$. The training algorithm is given as follows for the input signal $\mathbf{x}_{train}(t)$, a reservoir of size N_r and utilizes leaky integrator reservoir units:

1. Randomly generate the weight the matrices \mathbf{W}_{input} , \mathbf{W}_x at the input and reservoir layer.
2. Propagate the signal through the input and reservoir layers to construct a reservoir state vector for each $x_i(t)$ in the training set $\mathbf{x}_{train}(t)$ for $1 \leq i \leq m$. The reservoir state vector \mathbf{v}_i is an $N_r \times 1$ column vector initialized as $\mathbf{v}_1 = \mathbf{0}$ and is iterated as:

$$(20.4) \quad \mathbf{v}_{i+1} = (1 - \alpha)\mathbf{v}_i + \alpha f_{act}(\mathbf{W}_{input}x_i(t) + \mathbf{W}_x\mathbf{v}_i + \mathbf{W}_b)$$

for $i = 1, 2, \dots, m$ and $\alpha \in (0, 1)$ is the leaking rate which controls the speed of the reservoir dynamics in reaction to the input and \mathbf{W}_b is a randomly generated bias matrix. Furthermore, the quantity $f_{\text{act}}(\cdot)$ is taken as the activation function.

3. For every $i > \omega$, update the internal state of the network $\mathbf{X}_{(N_r+2) \times (m-\omega-1)}$

$$(20.5) \quad \mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_i(t) & x_{i+1}(t) & \cdots & x_m(t) \\ \mathbf{v}_i & \mathbf{v}_{i+1} & \cdots & \mathbf{v}_m \end{bmatrix}$$

where $\omega < m$ is the washout unit which is some initial transient to be discarded.

4. Define $\bar{\mathbf{x}}(t)$ to be the training set accounting for the discarded transient quantity ω . That is, $\bar{\mathbf{x}}(t)$ is taken to be:

$$\bar{\mathbf{x}}(t) = [x_{\omega+1}(t), x_{\omega+2}(t), \dots, x_m(t)]^T$$

The output matrix $\mathbf{W}_{\text{output}}$ can finally be found through a readout operation on $\bar{\mathbf{x}}(t)$ and the internal state \mathbf{X} :

$$\mathbf{W}_{\text{output}} = f_{\text{readout}}(\bar{\mathbf{x}}(t), \mathbf{X})$$

The readout operation is taken as a regression either in the form of the Tikhonov regularized regression (Ridge-Regression) or the Moore-Penrose inverse (Pseudo-Inverse):

$$\begin{aligned} f_{\text{readout}}(\bar{\mathbf{x}}, \mathbf{X}) &= \bar{\mathbf{x}}^* (\mathbf{X}^* \mathbf{X})^{-1} \mathbf{X}^* && \text{Pseudo-Inverse} \\ f_{\text{readout}}(\bar{\mathbf{x}}, \mathbf{X}) &= \bar{\mathbf{x}}^* (\mathbf{X}^* \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^* && \text{Ridge-Regression} \end{aligned}$$

for the regularization λ used to give preference to outputs having smaller norms. Once $\mathbf{W}_{\text{output}}$ is calculated, the network is effectively trained and ready to be validated. In order to validate the model, define the validating set $\mathbf{x}_{\text{val}}(t) = [x_{m+1}(t), x_{m+2}(t), \dots, x_n(t)]^T$ and repeat steps 2 and 3 from the training algorithm above for $\mathbf{x}_{\text{val}}(t)$ and construct the reservoir state vector and internal state. Once the internal state for the validation set has been calculated, the network's output $\mathbf{y}(t)$ can be found:

$$(20.6) \quad \mathbf{y}(t) = \mathbf{W}_{\text{output}} \mathbf{X}_{\text{val}}$$

where \mathbf{X}_{val} is the internal state that is constructed on the validation set $\mathbf{x}_{\text{val}}(t)$. In order to make out of sample estimates, the ESN is simply put into what is known as a 'free running' state. That is, given the network's validated internal state, the quantity $x_{n+1}(t)$ is estimated using the previous value $x_n(t)$. The ESN can be put in free run by utilizing $x_n(t)$ to calculate the reservoir state vector as:

$$(20.7) \quad \mathbf{v}_{n+1} = (1 - \alpha) \mathbf{v}_n + \alpha f_{\text{act}}(\mathbf{W}_{\text{input}} x_n(t) + \mathbf{W}_x \mathbf{v}_n + \mathbf{W}_b)$$

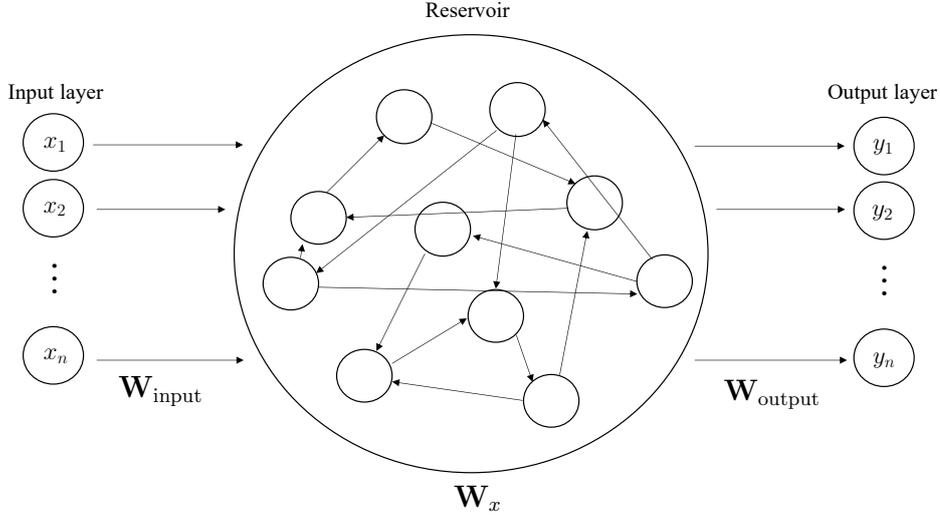


Figure 56: Schematic figure for the structure of an echo state neural network.

Once the reservoir state is calculated, the internal state is updated

$$(20.8) \quad \mathbf{X}_n = \begin{bmatrix} 1 \\ x_n(t) \\ \mathbf{v}_n \end{bmatrix}$$

and the successive quantity of interest is estimated as:

$$(20.9) \quad x_{n+1}(t) = \mathbf{W}_{\text{output}} \mathbf{X}_n$$

The robustness of the network can be cross validated by selecting a portion of data for testing that serves as a performance metric for the network's free run signal. Pathak et al. [60] demonstrated that low prediction error had been obtained for up to roughly 8 Lyapunov times (the average time for errors to grow by a factor of e) and can serve as a benchmark for the model's performance in this analysis. That is, upon the deployment, one seeks to maximize the total number of Lyapunov times that can resolve good performance, ideally seeking a value greater than 8.

20.4 Data Pre-processing

The pre-processing procedure of the input data begins with redefining the data in terms of a z -score metric which effectively standardizes the input. Furthermore, the data is assumed to be separable and can be decomposed as:

$$\dot{M}(t) = \dot{M}_d(t) + \dot{M}(t)$$

where $\dot{M}_d(t)$ is representative of the deterministic trend and $\dot{M}(t)$ is representative of the dynamics of the fluctuations. The fluctuating component $\dot{M}(t)$ is retrieved by inputting the

time series to a high pass filter having some passband frequency f_p with a given sampling rate f_s . Once $\dot{M}(t)$ is found, the deterministic component $\dot{M}_d(t)$ is easily retrieved. It is found that $\dot{M}_d(t)$ is typically not a stationary process and so it is common to utilize a d^{th} order difference operator B^d to induce some stationary behavior. The fluctuating and deterministic components are trained and validated independently on two separate networks. Each network is further placed in free run and the two signals are summed in order to produce the final output $\dot{M}(t)$.

20.5 Network Architecture

Echo state neural networks have been demonstrated to being highly elevated forms of recurrent models, but the generation of the random matrices $\mathbf{W}_{\text{input}}$ and \mathbf{W}_x can sometimes present a series of statistical challenges in simulating stable outputs. In certain circumstances, the randomness introduced by $\mathbf{W}_{\text{input}}$ and \mathbf{W}_x propagate into the model output resulting in several different (but similar) realizations of potential forecasts. Several alternatives to the traditional ESN have been demonstrated such as the Parallel ESN [54], the Deep ESN [61] and ensemble methods [55]. The Parallel ESN is an extension to its shallow counterpart where multiple reservoirs are stacked in parallel to one another. Thus, rather than having a single $\mathbf{W}_{\text{input}}$ and \mathbf{W}_x , the network will consist of M independently generated matrices $\mathbf{W}_{\text{input}}^{(j)}$ and $\mathbf{W}_x^{(j)}$ for $1 \leq j \leq M$. The input signal is trained and validated at each reservoir independently and a respective output matrix $\mathbf{W}_{\text{output}}^{(j)}$ is computed in accordance to the methodology outlined in section 20.3 for all reservoirs in the series. An output $\mathbf{y}^{(j)}(t)$ can be computed for each individual reservoir unit and the network's final output signal is taken as the mean of all individual paths generated. That is,

$$(20.10) \quad \hat{\mathbf{y}}(t) = \frac{1}{M} \sum_{j=1}^M \mathbf{y}^{(j)}(t)$$

The main drawback of the Parallel ESN is that there typically exists a path $\mathbf{y}^{(j)}(t)$ that outperforms the mean path $\hat{\mathbf{y}}(t)$. Given this observation, rather than taking a full average, one can perform a weighted average seeking an optimal set of weights $\hat{\beta}_j$ that minimizes the error of the final output signal. Therefore, the final signal is taken as a linear combinations of all realizations:

$$(20.11) \quad \hat{\mathbf{y}}(t; \hat{\beta}) = \sum_{j=1}^M \hat{\beta}_j \mathbf{y}^{(j)}(t)$$

where the optimal weights are found by treating this as a constraint optimization problem of the form:

$$(20.12) \quad \begin{aligned} & \min_{\hat{\beta}_j} \Delta(t; \hat{\beta}) \\ & \text{subject to } \sum_j \hat{\beta}_j = 1 \end{aligned}$$

for the sum of squared residuals

$$(20.13) \quad \Delta(t; \hat{\boldsymbol{\beta}}) = \left\| \mathbf{x}_{\text{opt}}(t) - \hat{\mathbf{y}}(t; \hat{\boldsymbol{\beta}}) \right\|^2$$

where $\mathbf{x}_{\text{opt}}(t)$ is the portion of the input used for optimization and cross validation of the free run signal. Furthermore, it is required that the error of the weighted output signal be at least less than or equivalent to the smallest error of any given path.

$$(20.14) \quad \text{error}(\hat{\mathbf{y}}(t; \hat{\boldsymbol{\beta}})) \leq \text{error}(\mathbf{y}^{(j)}(t)) \quad \forall j \in (1, M)$$

The critical assumption in this methodology is that the set of weights that optimize the free run segment will additionally produce the best performing output outside of this regime. Therefore, in order to implement the optimal weighted strategy for out of sample estimates, one implements the following algorithm:

1. Split the data up into the respective training, validation and free running segments:
 - Training: $t_0 < t \leq t_{\text{tr}}$
 - Validating: $t_{\text{tr}} < t \leq t_{\text{vl}}$
 - Free run: $t_{\text{vl}} < t \leq T$
2. Randomly generate M input and reservoir matrices independantly.
3. Train, validate and compute the respective output matrices $\mathbf{W}_{\text{output}}^{(j)}$ for each reservoir in the series according to the methodology discussed in section 20.3.
4. Simulate M independant free running signals $\mathbf{y}^{(j)}(t)$ from each reservoir up to some $t > T$.
5. Concatenate each realization into a path matrix $\mathbf{Y}_{n \times M}$ defined as:

$$(20.15) \quad \mathbf{Y} = \begin{bmatrix} y_1^{(1)} & y_1^{(2)} & \cdots & y_1^{(M)} \\ y_2^{(1)} & y_2^{(2)} & \cdots & y_2^{(M)} \\ \vdots & \vdots & \ddots & \vdots \\ y_n^{(1)} & y_n^{(2)} & \cdots & y_n^{(M)} \end{bmatrix}$$

6. Solve the constrained linear optimization problem defined by 20.12 on the free run segment $t_{\text{vl}} < t \leq T$ of the all realizations for the set of weights $\hat{\beta}_j$
7. Compute the weighted sum of all free running signals for $t > T$ using equation 20.11.

The column vector $\hat{\boldsymbol{\beta}}$ will have length equal to the total number of reservoirs M in the series. A schematic diagram on the network's architecture is demonstrated in figure 57. Note, the network's optimized output will only be as good as the performance of the individual free running signals $y^{(j)}(t)$. Therefore, if each output path has poor performance, the network's final optimized output will additionally also be poor.

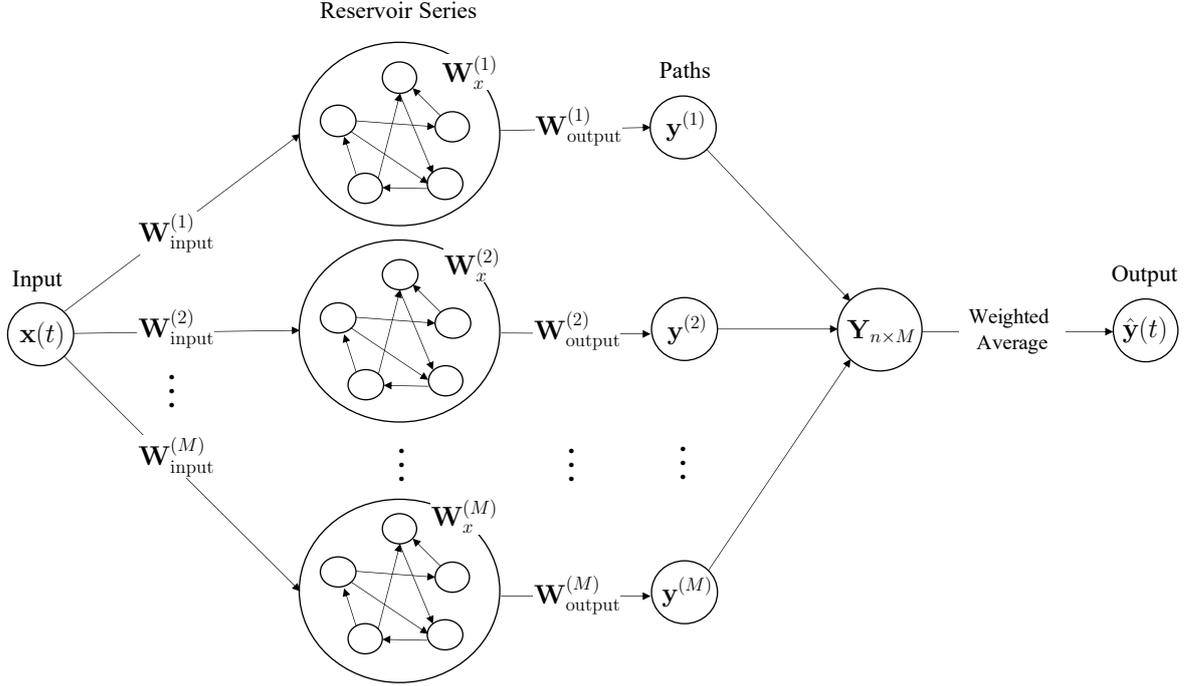


Figure 57: Schematic figure for the architecture of a parallel ESN utilizing optimal weighted averaging in extracting the final free running network output.

20.6 Hyperparameter Optimization

Hyperparameter tuning is crucial in instantiating a robust and efficient reservoir. A popular method in hyperparameter optimization is the grid search method and its popularity lies in its simplicity. Grid search effectively deploys an exhaustive searching protocol through a specified subset of the hyperparameter space. In performing such an optimization, a range of hyperparameters are defined for the search and summarized below.

Table 20.2: Search Space for the grid search optimization.

Parameter	Search Space
Reservoir Size, N_r	(250, 1000) incremented by 250
Spectral Radius, ρ_s	(0.1, 1) incremented by 0.1
Input Scaling, ρ	(0.1, 1) incremented by 0.1
Leaking Rate, α	(0.1, 1) incremented by 0.1
Connectivity, \tilde{C}	[0.01, 0.05, 0.1]

Additionally, the deployed grid search also surveys various architectures for reservoir sampling. That is, since the sparse matrices $\mathbf{W}_{\text{input}}$ and \mathbf{W}_x are generated randomly, they can be constructed through different methods of random sampling. The architectures studied will be reservoirs that are uniformly generated, normally generated and lognormally generated. Table

20.2 demonstrates a search for reservoir sizes in the range 250 to 1000; a preliminary survey has also been implemented on reservoir sizes up to and including 5000 and have shown to give little to no improvement in the network's final output thus have been omitted from the final grid search routine. In order to satisfy the Echo State Property (ESP), the spectral radius is bounded to $0.1 \leq \rho_s \leq 1$. The ESP is given as a property that the state of the network will asymptotically depend only on the driven input signal and that dependencies on the initial conditions will eventually near zero. An echo state neural network is said to satisfy the ESP when $\forall \mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)] \in (\mathbb{R}^{N_{\text{in}}})^n$ and $\forall \mathbf{v}_0, \mathbf{v}'_0 \in \mathbb{R}^{N_R}$ initial states:

$$(20.16) \quad \|\mathcal{F}(\mathbf{x}, \mathbf{v}_0) - \mathcal{F}(\mathbf{x}, \mathbf{v}'_0)\| \rightarrow 0 \quad \text{for } n \rightarrow \infty$$

where

$$\mathcal{F}(\mathbf{x}, \mathbf{v}_0) = (1 - \alpha)\mathbf{v}_0 + \alpha f_{\text{act}}(\mathbf{W}_{\text{input}}\mathbf{x}(t) + \mathbf{W}_x\mathbf{v}_0 + \mathbf{W}_b)$$

is the state transition function. The spectral radius is taken as the absolute maximum eigenvalue of the reservoir weight matrix:

$$(20.17) \quad \hat{\rho}_s(\mathbf{W}_x) = \max(|\text{eig}\{\mathbf{W}_x\}|)$$

and so it is found that scaling the reservoir weight matrix can guarantee that the ESP is satisfied. The ESP condition is sufficiently met if the input signal demonstrates contractive dynamics, that is if:

$$\|\mathbf{W}_x\|_2 < 1$$

which is satisfied for $\hat{\rho}_s(\mathbf{W}_x) < 1$. Thus, when generating the reservoir weight matrix, the initialization is given as:

$$(20.18) \quad \mathbf{W}_x = \tilde{\mathbf{W}} \frac{\rho_s}{\hat{\rho}_s(\tilde{\mathbf{W}})}$$

where $\tilde{\mathbf{W}}$ is generated at random and ρ_s is the desired spectral radius taken as a hyperparameter. The input scaling ρ controls the amount of non-linearity in the dynamics of the reservoir. Larger input scaling will invoke a more nonlinear response in reservoir dynamics. Lastly, the connectivity \tilde{C} controls the sparsity of the reservoir weight matrix and is taken to illustrate the degree of connections amongst the units in the network's reservoir. That is, if a reservoir is said to have 1% connectivity, the weight matrix is considered to be highly sparse. The search space of the grid search algorithm will also be extended to include a regularization parameter if the readout operation utilizes Ridge-Regression. Thus, the first search will deploy the Moore-Penrose Pseudo-Inverse and the second search will implement Ridge-Regression, including a search over the regularization parameter $\lambda \in [1, 10, 100, 1000] \times 10^{-8}$. The optimal hyperparameter outputs are summarized in table 20.3 where the grid search was performed on the validation set using a single reservoir on the fluctuating component of the data. Rather than performing a separate search on the deterministic component, the same hyperparameters are used for both $\dot{M}_d(t)$ and $\dot{M}(t)$. In all cases, an input bias scaling of 0.1 is used with $f_{\text{act}}(x) = \tanh(x)$.

Table 20.3: The optimal hyperparameters for all models.

Model	Hyperparameters						Readout	Sampling	MSE
	N_r	ρ_s	ϱ	α	\tilde{C}	λ			
Model 26	250	1.0	0.5	1.0	0.05	-	Pseudo	Uniform	0.0078
	250	0.9	0.1	0.1	0.01	10^{-6}	Ridge	Uniform	0.0028
Model 27	250	1.0	1.0	1.0	0.05	-	Pseudo	Uniform	0.0111
	250	0.7	0.1	0.2	0.01	10^{-5}	Ridge	Uniform	0.0019
Model 28	250	1.0	1.0	1.0	0.05	-	Pseudo	Uniform	0.1693
	250	0.2	0.1	0.1	0.05	10^{-5}	Ridge	Uniform	0.1527
Model 29	250	1.0	0.1	0.2	0.01	-	Pseudo	Normal	0.0033
	500	0.9	0.2	0.4	0.05	10^{-7}	Ridge	Uniform	0.0028
Model 30	250	1.0	1.0	1.0	0.05	-	Pseudo	Uniform	0.0046
	250	1.0	0.1	0.3	0.1	10^{-5}	Ridge	Uniform	2.52×10^{-4}
Model 31	250	0.8	0.1	0.4	0.01	-	Pseudo	Uniform	0.0072
	500	0.9	0.1	0.3	0.05	10^{-7}	Ridge	Uniform	0.0068
Model 32	500	0.1	0.1	0.2	0.01	-	Pseudo	Uniform	3.32×10^{-4}
	250	1.0	0.1	0.4	0.1	10^{-7}	Ridge	Uniform	2.67×10^{-4}

20.7 Predictive Analysis

Training and validation are performed on each model individually and deployed only on the chaotic regime of the time series. Noting that the simulation data for models 26 to 32 incorporate adaptive time steps, the data has been mapped onto an evenly spaced time domain in an effort to improve visualization of the accretion dynamics. The selection of the training, validation and free run segments aimed to capture as much of the chaotic portions of the model as possible without incorporating too much of the flat lined dynamics exhibited at later periods. Given that certain models exhibit longer episodic periods than others, the number of data points available for training and validation vary from one model to the next. In all cases, 50 data points will be used in the free running segment and will be given in terms of the respective maximum Lyapunov exponent. The data is decomposed into a deterministic and fluctuating component as discussed in section 20.4 where each component will be free run separately and summed to demonstrate the model's final output. The network implements $M = 100$ reservoirs in the series and additionally extends to studying response to higher non-linearity in the reservoir dynamics by increasing the respective input scaling ϱ . That is, if the input scaling is increased, the model tends to exhibit higher variability in each realization from the reservoir series. It can be demonstrated that having high variability in the outputs can sometimes be advantageous for optimal weighted averaging as it can allow the network to capture more intrinsic dynamics of the target output. The analysis below pertains only to model 29 of the simulation data, but the findings and outputs for all remaining models can be found in appendix A1.6. Model 29 had a chaotic regime containing roughly 10050 data points where 5000 were used for training, 5000 for

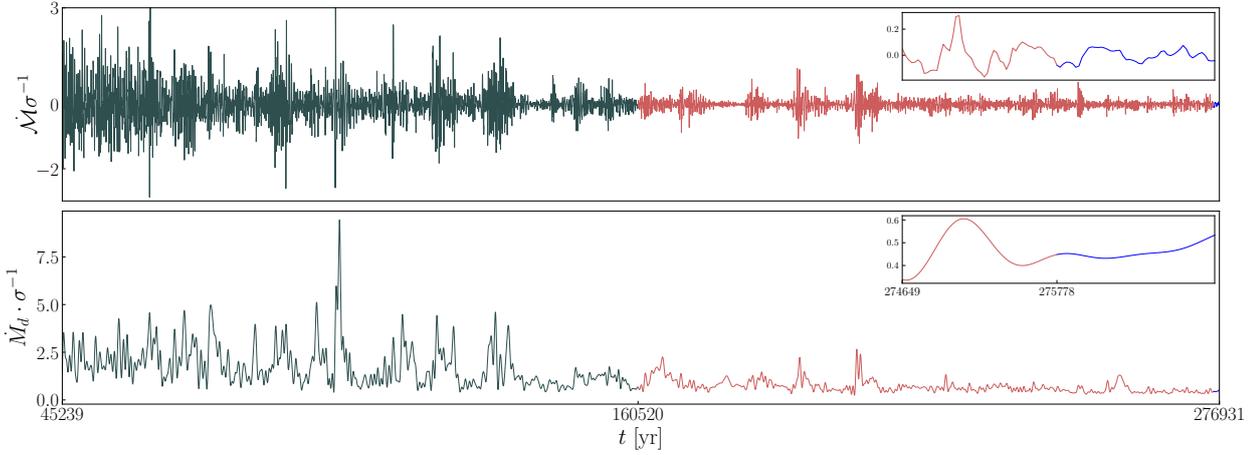


Figure 58: The training/validation/free running segments for model 29. **Top:** Demonstration of the standardized fluctuations in the model’s dynamics. **Bottom:** The model’s standardized deterministic dynamics found by subtracting the fluctuating component from $\dot{M}(t)$.

validating and 50 for free run which translate to roughly 9.2 Lyapunov times. The standardized data segments for the deterministic and fluctuating components are given in figure 58 where the model’s fluctuating components were extracted by implementing an elliptic high pass filter having a passband frequency $f_p = 2000$ Hz and a sampling rate $f_s = 50000$ Hz. The respective hyperparameters are adopted from table 20.3 including a washout quantity $\omega = 100$. The respective free running signals for $\dot{M}(t)$ and $\dot{M}_d(t)$ adopting the optimal hyperparameters are demonstrated in the top two plots for panels (a) and (b) from figure 59 respectfully. Imposing an increased non-linear response in the network’s reservoir by increasing the input scaling to $\varrho = 1$ gives rise to the outputs in the bottom two plots of panels (a) and (b) in figure 59. Summing the individual components in the respective plots in panels (a) and (b) in figure 59, the final, non-standardized output is given in figure 60. Table 20.4 additionally provides a summary the respective performance metrics on each component of model 29 in terms of the mean squared

Table 20.4: Summary of the respective performance metrics for model 29.

Low input scaling			High input scaling ($\varrho = 1$)		
Readout	MSE	Component	Readout	MSE	Component
Ridge	0.00069	$\dot{M}(t) \cdot \sigma^{-1}$	Ridge	0.00048	$\dot{M}(t) \cdot \sigma^{-1}$
Pseudo	0.00051	$\dot{M}(t) \cdot \sigma^{-1}$	Pseudo	0.00024	$\dot{M}(t) \cdot \sigma^{-1}$
Ridge	0.00017	$\dot{M}_d(t) \cdot \sigma^{-1}$	Ridge	1.75460e−5	$\dot{M}_d(t) \cdot \sigma^{-1}$
Pseudo	5.30510e−6	$\dot{M}_d(t) \cdot \sigma^{-1}$	Pseudo	4.76969e−6	$\dot{M}_d(t) \cdot \sigma^{-1}$
Ridge	0.00089	$\dot{M}(t) \cdot \sigma^{-1}$	Ridge	0.00049	$\dot{M}(t) \cdot \sigma^{-1}$
Pseudo	0.00049	$\dot{M}(t) \cdot \sigma^{-1}$	Pseudo	0.00025	$\dot{M}(t) \cdot \sigma^{-1}$

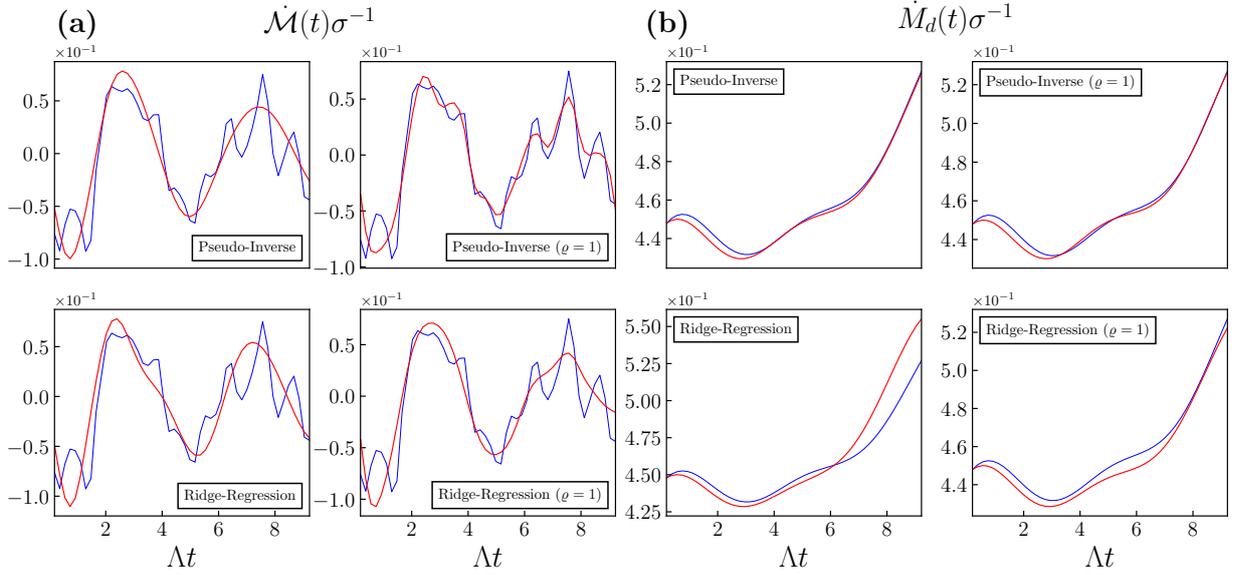


Figure 59: The optimized free running signal (red) utilizing a series of $M = 100$ reservoirs overlaid with the free run segment (blue) of both the standardized (a) fluctuating component and (b) deterministic component. Each output is given in terms of the model's respective maximum Lyapunov exponent.

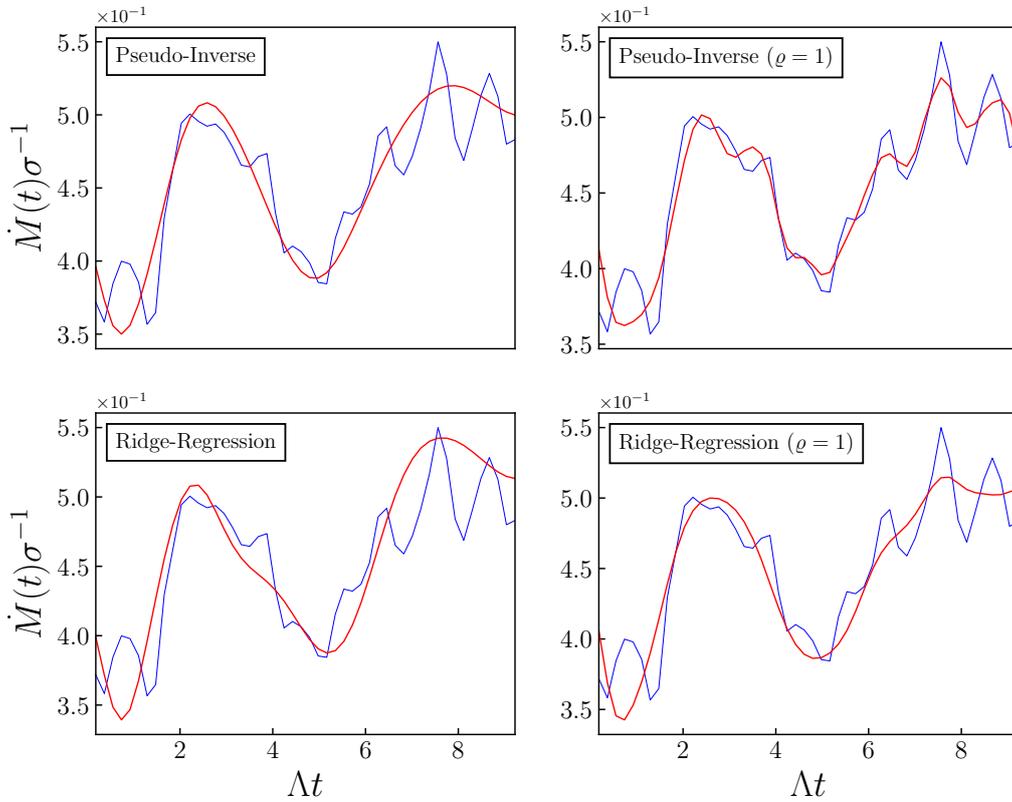


Figure 60: The network's final optimized output signal utilizing a series of $M = 100$ reservoirs found by summing the free running output for the deterministic and fluctuating components overlaid with the free run segment.

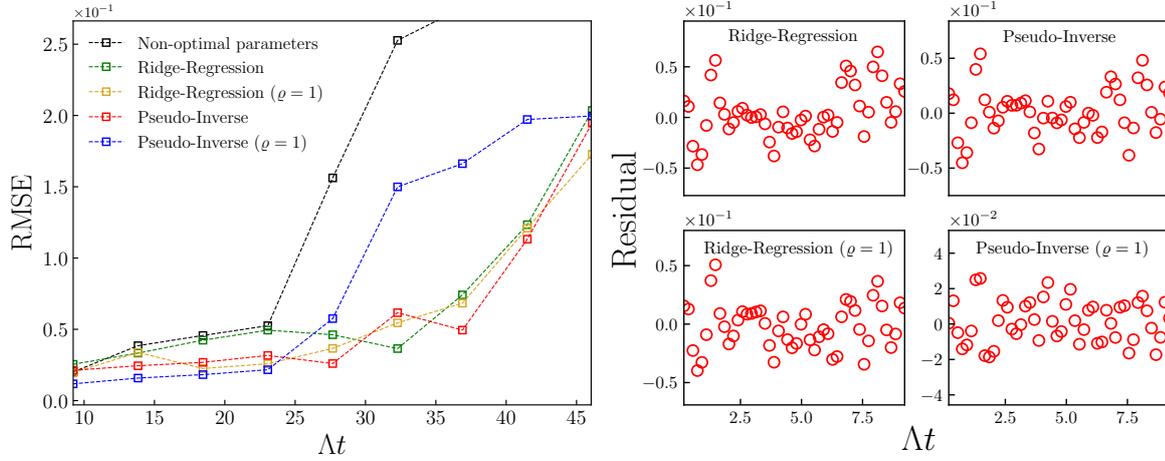


Figure 61: The time evolving RMSE demonstrated for $\dot{M}(t) \cdot \sigma^{-1}$ up to 46.1 Lyapunov times (**left**) and the model residuals up to 9.2 Lyapunov times (**right**).

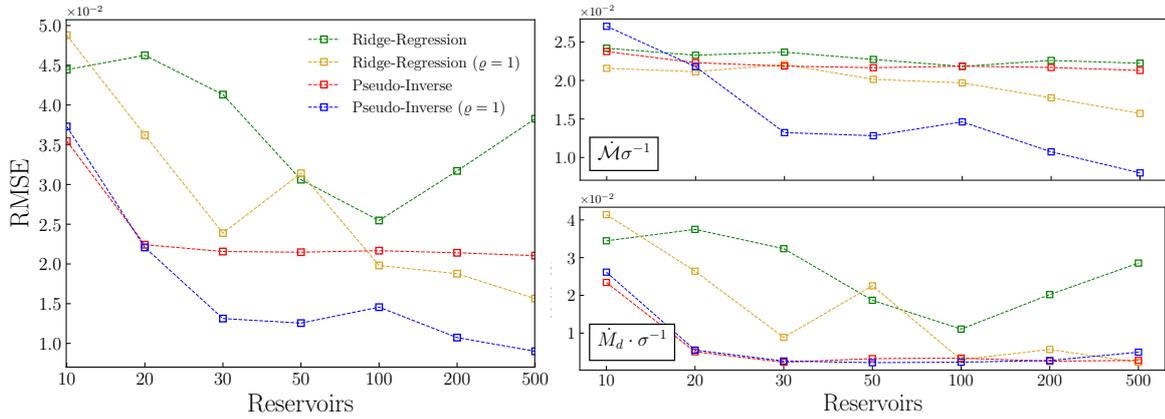


Figure 62: The RMSE demonstrated as a function of the number of reservoirs in the series. The right panels demonstrate the same respective RMSE for the individual fluctuating (**top**) and deterministic (**bottom**) components.

error. In order to quantify the network’s long term performance, the final output signal was additionally run for 250 time steps (≈ 46 Lyapunov times), recording the RMSE after every 25 time steps, demonstrated in the left panel of figure 61. Additionally, a series of residual plots are given in the right panel of figure 61 for all respective output signals. In order to demonstrate the sensitivity in the network’s performance with regard to the respective hyperparameters, an additional set of hyperparameters deemed non-optimal are demonstrated by the black line in the left panel of figure 61. The respective non-optimal hyperparameters take $N_r = 100$, $\alpha = 0.4$, $\rho_s = 0.1$, $\varrho = 0.5$ and $\tilde{C} = 0.25$ with lognormally sampled input and reservoir matrices implementing the Pseudo-Inverse readout. This network is well out of the chaotic state demonstrated in the optimal models due to its low spectral radius. For short term outputs up to approximately $\Lambda t \approx 23$, all networks demonstrate similar performance. Beyond this point, the non-optimal network errors become excessively amplified. In the optimal network models,

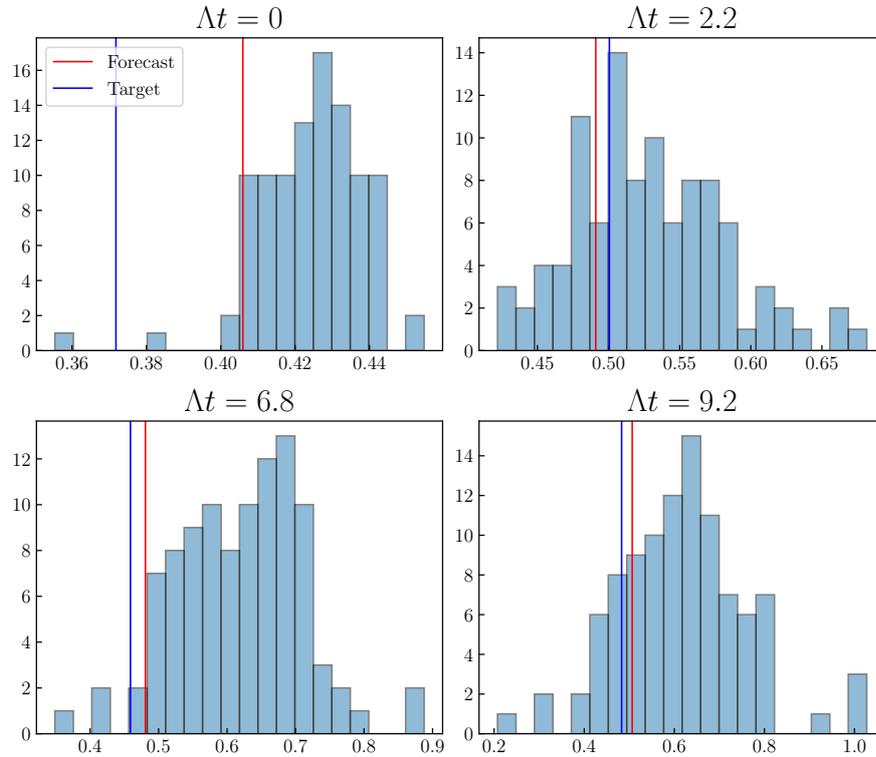


Figure 63: The distribution of outputs from the reservoir series at four instances in time implementing a network series of $M = 100$ reservoirs.

good performance is demonstrated in all models up to roughly 28 Lyapunov times. Beyond this point, the RMSE demonstrates exponential growth. A discussion can be made on whether there is any direct correlation between performance and the total number of reservoirs in a series. Figure 62 shows how the RMSE changes with total number of reservoirs. The left panel demonstrates the RMSE for the standardized final output $\dot{M}(t) \cdot \sigma^{-1}$ and the right panel figures are for the individual fluctuating and deterministic components. Upon visual inspection, there appears to be significant improvement up to and including 30 reservoirs with a potential plateau beyond this point. Whether there is truly a plateau or not raises the question on scaling and whether there is additional improvement at any point beyond $M > 500$ reservoirs. For large enough number of reservoirs, computational efficiency can become significantly compromised, but given that training and validation at each reservoir are independent, it becomes natural to deploy the OWA-ESN in a parallel environment.

20.8 Prediction Bands

The framework that grounds the OWA-ESN involves the simulation of M free running signals where M is the total number of reservoirs in the network series. Given this, at any given point in time t_i , there exists a distribution of M possible estimates $y_i^{(j)}$ for $1 \leq j \leq M$. Figure

63, demonstrates the output distributions at four different instances in time for model 29 implementing a network using a ridge-regression readout adopting the hyperparameters from table 20.3 and incorporating a high input scaling ($\varrho = 1$). The distributions are shown to not always be perfectly Gaussian and the OWA-ESN optimized output is not always governed by the bin with the largest count. Additionally, these distributions can be used to construct a band of estimates at each time point by computing a q^{th} interpercentile range of the distribution. Figure 64 demonstrates how prediction bands (shaded red region) are incorporated into the standardized outputs for the model's fluctuating and deterministic components. Furthermore, the components can again be summed to demonstrate the network's final output given in figure 65. This figure is identical to the bottom left panel of figure 60, but now includes the network's prediction band found by computing a 99% interpercentile range at each time step.

20.9 Robustness & Hybrid Networks

The critical assumption for the OWA-ESN is that if the optimized output signal is run beyond the free running segment, it will remain the optimal output. Furthermore, the networks deployed thus far have been the same for both the deterministic and fluctuating components. That is, both the deterministic and fluctuating components of the input data were fed into the same reservoir series. Given that these components exhibit different temporal dynamics, a hybrid network can be implemented to exploit these differences. Figure 66 schematically demonstrates the hybrid network and the flow of data from the input $\dot{M}(t)$ to an estimated output $Y(t)$. Each square panel captioned with "Reservoir Series" is representative of the parallel ESN with weighted averaging demonstrated in figure 57. The hybrid network can benefit from two reservoir series with different hyperparameters used to optimize the respective input component. In doing such, a hybrid OWA-ESN is deployed on model 29 of the simulation data and now allows for an additional 50 time points outside the free run segment. The hyperparameters are adopted from table 20.5 where the random reservoir initialization is seeded for reproducibility with a start

Table 20.5: Hybrid network hyperparameters.

Component	N_r	ρ_s	ϱ	α	\tilde{C}	λ	Readout	Sampling
Fluctuating	250	1.0	1.0	0.2	0.01	-	Pseudo	Normal
Deterministic	250	1.0	0.1	0.2	0.01	-	Pseudo	Normal

value $s_0 = 503489$ and iterated throughout the series in accordance to the linear congruential generator (LCG) algorithm,

$$(20.19) \quad s_{n+1} = (as_n + b) \bmod c$$

for the coefficients $a = 2^{16} + 3$, $b = 0$ and $c = 2^{31}$. Here, the recurrence relation is run for a total $n = 1, 2, \dots, M$ where the same sequence is used for both reservoir series. The network's

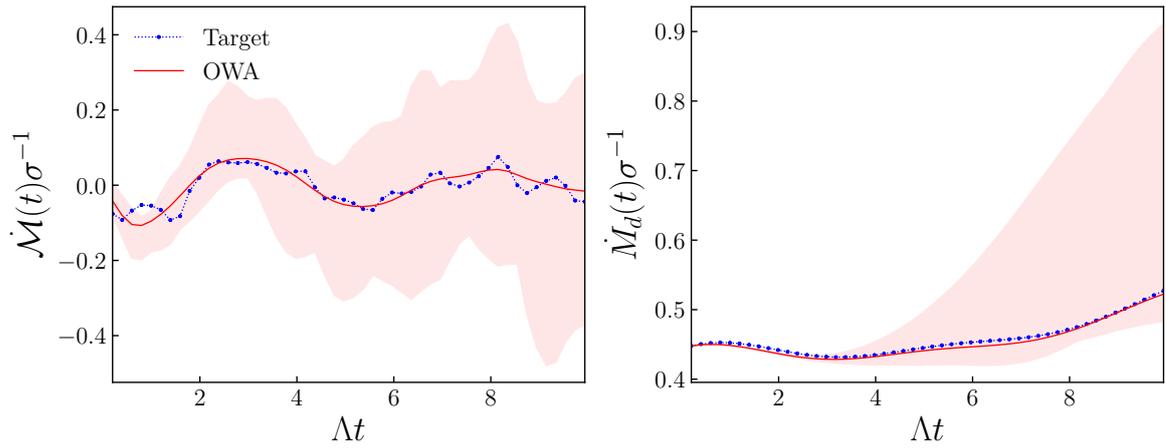


Figure 64: The optimized output signal including the prediction bands for model 29 where the left plot is demonstrative of the model’s fluctuating component and the right plot is the model’s deterministic component. The red shaded region represents the 99th interpercentile measurement of the distribution at all time points t_i .

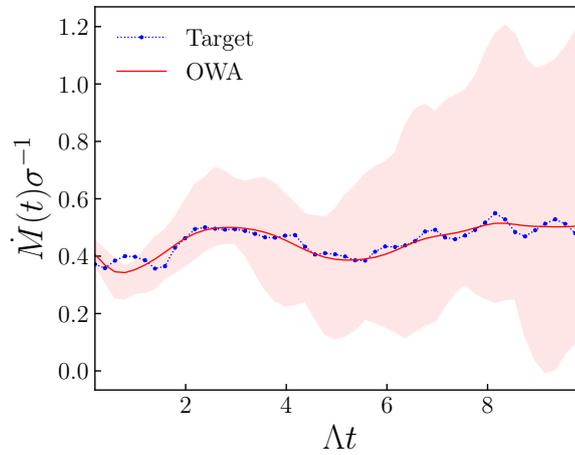


Figure 65: The network’s optimized final output incorporating prediction bands found by summing the two components in figure 64. The red shaded region represents the 99th interpercentile measurement of the distribution at all time points t_i .

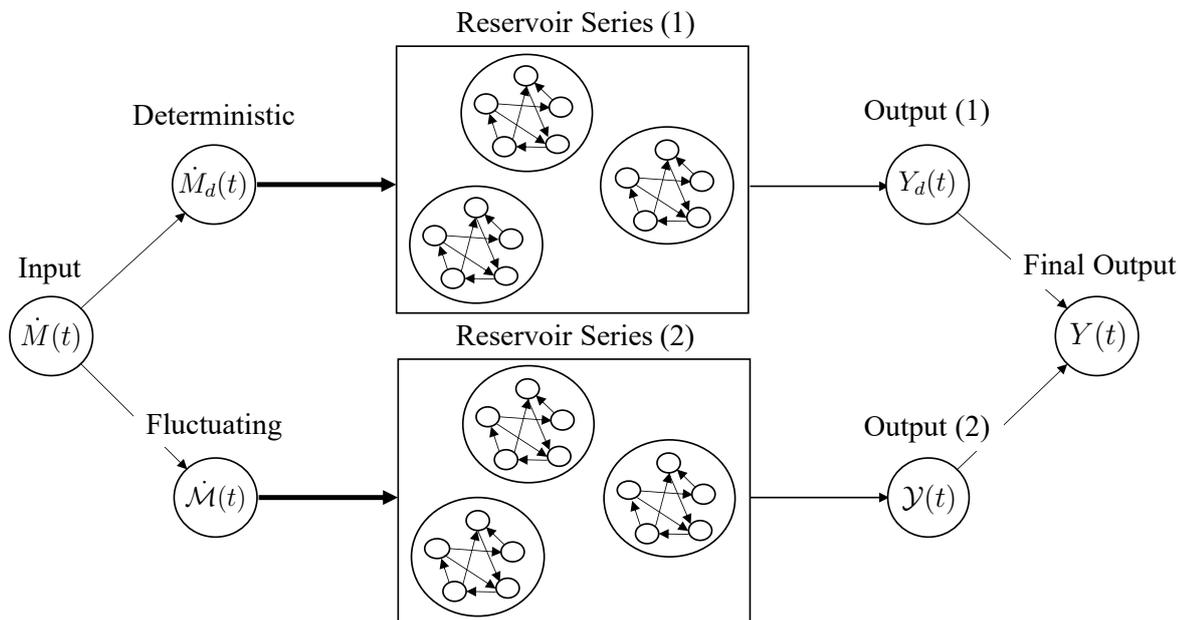


Figure 66: Schematic figure demonstrating the implementation of the hybrid network. After pre-processing the input data, each component is fed through two separate reservoir series each optimized differently. Finally, the network's final output is summed.

robustness is back tested by letting the hybrid network free run for a total of 100 time points where the first 50 will be used to optimize the signal. The deployment utilizes $M = 100$ reservoirs in each series and the training and validation segments will be the same as that in section 20.7. In figure 67, the outputs for both the fluctuating and deterministic components are given for the network's optimized signal. Additionally, the black line is representative of the network's top performing individual signal. That is, of the 100 individual paths given by the reservoir series, the top performing output is taken as the one with the lowest MSE in the free run segment. The constraints of OWA-ESN require the OWA output be at least as good as the top performing one in the free run segment. The final outputs are given in figure 68 where the MSE of the hybrid OWA signal is 0.00085 and that of the best individual signal is 0.00114. Visually inspecting the components in figure 67, it can be shown that the best individual path is comparable for the deterministic component, but rapidly decays and flat lines outside the free running segment for the fluctuating component. The optimized output however, captures the important temporal features exhibited in both the deterministic and fluctuations in the data. In the left panel of figure 67, the OWA output exhibits similar oscillations outside the free running segment to that of the target output. Although the oscillations become more out of phase as the output progresses further out of the free running segment, the intrinsic dynamics of the data appear to have been extracted by the network to some degree and demonstrate an improvement over the network's best individual output. Similarly in the right panel of figure 67 which is representative of the deterministic trend, the optimized signal appears to have additionally

extracted the increasing behavior of the output outside the free run segment. When the two components are summed to produce the final output, again it is shown that the optimized output signal appears to have extracted the important temporal features of the simulation data along with the general trend and behavior of the model dynamics. These dynamical properties are not demonstrated for the longer term behavior of the individual best performing signal. Furthermore, in the right panel of figure 68, the hybrid output is compared to implementing the non-hybrid OWA-ESN where the input scaling is held at $\rho = 0.1$.

The OWA-ESN aimed to exploit some of the statistical challenges of the shallow ESN. The random generation of the input and reservoir matrices in shallow ESNs can introduce stochastic nature to the outputs which can be cumbersome in practice. The weighted average approach aims to suppress some of the lower performing outputs while emphasizing the better performing ones within a given set of constraints. Over the free running segment, the optimized signal is indeed found to significantly mitigate the randomness in outputs. However, outside of the free run segment, the sensitivity to the network’s hyperparameters becomes more important. Particularly, models with high input scaling parameter introduce higher variability in the output. These models should be seeded for reproducibility and should be done so prior to any hyperparameter optimization. This will ensure that any given network is optimized relative to the given seed. This becomes less important when the input scaling is low. It was also found that high input scaling models exhibit stability issues for long range forecasts ($\gtrsim 280$ time steps, but can vary depending on the model’s other hyperparameters). This is due to the fact that a high input scaling results in large variability in the outputs which tend to diverge more rapidly. Deterministic approaches have been used to completely remove the stochastic nature of ESN type neural networks. Dan et al. 2014 [62] demonstrated an approach to generating fully deterministic input and reservoir matrices. They propose a Delay-Line reservoir topology where non-zero values are located on the lower sub-diagonal of the reservoir matrix having feed-forward weights given by a tuning parameter r . Additionally, the input layer is fully connected with weight value $|\nu| > 0$, where the sign is determined by generating a pseudo-random aperiodic sequence. This methodology was found to have comparable performance relative to the shallow ESN on benchmark chaotic time series such as the sunspot or NARMA data sets [62, 63]. The deterministic ESN counterpart can be incorporated with the optimal weighted averaging strategy where rather than seeding a random number generator, the tuning parameters r and ν are changed M times to generate different sets of deterministic reservoir topologies.

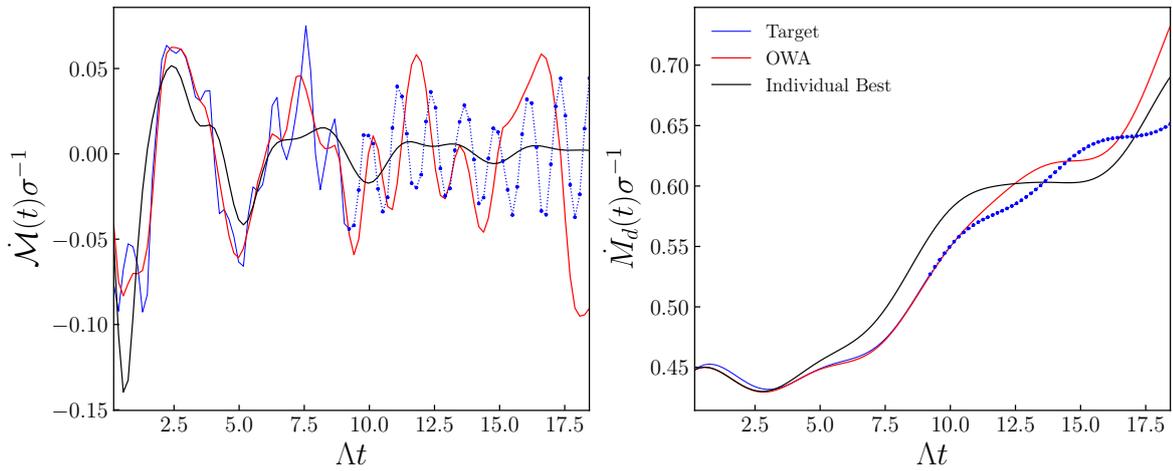


Figure 67: The optimized output signal along with the best individual output for both the fluctuating and deterministic components of the data run out of the free running segment. The solid blue line is representative of the portion of data used in the optimization, whereas the dotted dashed line is the remaining segment that is out of sample.

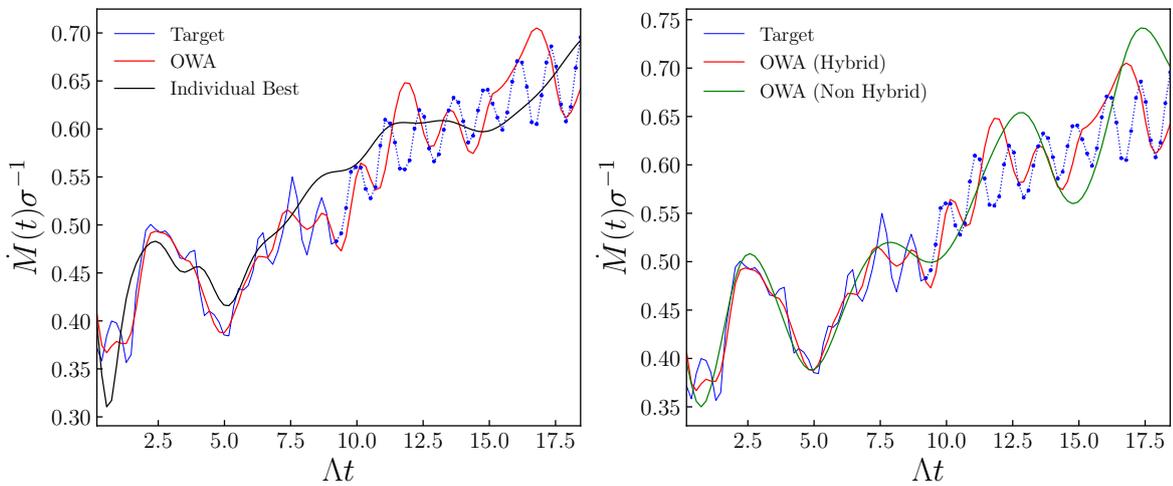


Figure 68: The optimized output signal along with the best individual output for the network's final output found by summing both components in figure 67. The solid blue line is representative of the portion of data used in the optimization, whereas the dotted dashed line is the remaining segment that is out of sample. **Left:** Comparison between the output from the hybrid network (red) with the individual best (black). **Right:** Comparison between the output from the hybrid network (red) with a non-hybrid network (green).

21 Applications, Conclusion & Next Steps

Time series machine learning and statistical models are commonly used to infer and extrapolate historical data out into the future to forecast temporal dynamics for systems that do not have any underlying model to describe its behavior. The OWA-ESN thus far has been applied to simulation data aiming to mimic the complex, chaotic dynamics of mass accretion in the late phase of protostellar evolution. Although many of these episodic systems are commonly modelled through some form of hydrodynamic simulations [14, 56, 57, 64], there are still phenomenon in star formation that are not well understood [65]. FU Orionis (FUor) are examples of such objects whose relationship to standard pre-main sequence evolution remains uncertain.

21.1 FUor & EXor Objects

FUor objects are a type of pre-main sequence stars which demonstrate rapid changes in magnitude and accretion that remain high over several decades. Ex Lupi (EXor) objects share a similar sequence of outbursts but over a much shorter time scale with lower accretion rates. The bursts associated with FUor objects are longer and stronger than EXors, occur cyclically and are driven by excess material from circumstellar envelopes. EXors are believed to be associated with instabilities in the disks of T-Tauri stars whereas FUors are associated with both protostars having a disk and envelope along with T-Tauri stars with disks [65]. However, observations have given light to a much more complex picture where it is unsure if these two classes of pre-main sequence objects are distinct from one another or whether there is a relationship between the two. Particularly, the outburst decay time and outburst repetition time differences between FUor and EXor classes remains in question. Given that FUor objects are believed to exhibit long duration eruptions, acquiring observational data to assess such dynamics becomes challenging given the natural time constraint. Theoretical models predict that FUor outbursts occur multiple times in the phase of early stellar evolution. Current mathematical models lack in some respect to modelling such objects completely. These include numerical limitations, inconsistencies in modelling the evolution of the entire disk and the lack of theoretical work associated with EXors [65]. The introduction of neural networks to such scenarios can significantly help in extracting and learning important temporal characteristics of such objects. Given small computational overhead and the proven ability to extract and extrapolate temporal dynamics associated with chaotic systems, neural networks can be used to forecast light curves and accretion bursts to give insight to some important characteristics for both FUor and EXor objects.

21.2 Caveats & Remedies

Although the OWA-ESN has demonstrated strong performance on the mass accretion simulation data, it is not free of any caveats that must be discussed. In order to truly reap the benefits of the optimal weighted average approach, appropriate hyperparameter tuning is crucial. Thus far,

the method of choice for hyperparameter optimization has been the grid search method. Given the size of the ESN hyperparameter space, grid search can be rather overkill and extremely time consuming. The grid search used in section 20.6 pertained only to the fluctuating component where the same hyperparameter outputs were used to model the deterministic component. This meant that although the OWA-ESN had still demonstrated promising results, the network had still not been utilized to its full potential. Had a secondary grid search been performed separate to the deterministic component, this would have added to an already computationally expensive search. Incorporating more efficient hyperparameter optimizations can not only expedite this process but can potentially lead to better outputs if the optimization is not constrained within a given search space. Given that the loss function is differentiable, one can seek to implement a stable gradient based method to optimize the network hyperparameters. Thiede & Parlitz 2019 [66] developed a stable gradient based optimizer that demonstrates comparable performance metrics to traditional grid search but has much improved speed.

Another important component of the OWA-ESN that must be addressed is how to properly seed the random number generator for the reservoir series. Note, the OWA-ESN exploits variability in output and so the initial seed given must be iterated in some fashion through the reservoir series as done in section 20.9. The method of iteration is not set in stone and can be done so however the model user chooses. It was found that the selection of the initial seed becomes less important when the input scaling is low and becomes very important for reproducibility when the input scaling approaches 1. To err on the side of caution, it is recommended to seed the generator prior to hyperparameter tuning.

Given the repetitive nature of the OWA-ESN in generating a distribution of outputs, the model can become computationally taxing if there are a large number of reservoirs in the series and if the size of input data is large. The most obvious method of overcoming this is to implement the network with parallel computing. Simply divide the number of reservoirs in the series evenly over the number of cores on a machine to generate each free run signal. The master processor can then perform the optimization after receiving the outputs from all cores.

21.3 Next Steps

Subsection 21.2 outlined some important developments that can and should be made to the OWA-ESN to improve both performance and accuracy. In addition to this, future work can look to extend network training to several data sets rather than one. That is, the OWA-ESN can be trained using the data from models 26 to 31 and can be validated and optimized on model 32 for forecasts. One aim is to use this methodology to predict when disk formation occurs during the core's evolution. At the time of disk formation the time series exhibits a transition from smooth to highly chaotic dynamics. This direction would train the network on a series of simulation models where the network would learn to estimate the approximate moment the dynamics are expected to shift from smooth to chaotic, indicative of disk formation.

CONCLUSION

This thesis aimed to study the evolutionary properties that govern star formation from the early prestellar phase up to protostellar development. Through the use of both analytic and numerical methods, one can provide a fully comprehensive study on many of the governing properties surrounding star formation. Beginning with prestellar evolution, the analytic magnetic field model first introduced by Ewertowski and Basu [9] proved to be highly effective in modelling the magnetic state of the observed core FeSt 1-457. The model robustness was back tested via the POLARIS code and enabled for a detailed discussion and inference on many properties governing the core's current and future states. The thesis then extended to the early phases of protostellar evolution which required more complex numerical models for accurate representation. Primarily, numerical MHD models served as a foundation of what became an extensive radiative transfer simulation performed in POLARIS. The outputs to these simulations provided high resolution maps that governed the polarization surrounding the small scale regions in protostellar cores. In doing so, the synthetic polarization maps allowed for the extrapolation and study of the magnetic field morphology along with the emission & absorption properties governing observed polarimetry data of the Orion Source I cloud. Finally, the late protostellar phase provided the framework for an introduction as to how neural network models can be used as a robust predictor for time series. A new form of recurrent neural networks were developed and demonstrated promising preliminary results in modelling the chaotic state surrounding the episodic accretion involved in late protostellar evolution. Such models are aimed to being an effective tool for data that lacks an effective underlying dynamical system model.

Each chapter in this thesis can be extended in future work. In chapter 2, the model did not include the possibility for a tilt angle of the core axis with the plane of the sky. Future modelling can incorporate this by fitting the emergent synthetic polarization maps with the observed maps. This would allow a better inference of the fully three-dimensional structure characterizing the magnetic field. Furthermore, a more elaborate fit can be made by using Bayesian statistical inference which involves computing a posterior distribution of values for each parameter, typically found through Markov Chain Monte Carlo. This approach allows for a distribution of potential fits which benefits from employing an entire grid of models. The MHD simulations in chapter 3 introduced an emergent magnetic field morphology involving a series of twists and irregularities that were not able to be modelled by the Ewertowski and Basu

[9] magnetic field functions. This is primarily due to the assumption of axisymmetry which allows one to project a single plane cut about the azimuthal angle. After the formation of a protostar, there is also significant twisting in the field lines that causes a pressure gradient which drives bipolar jets and outflows. Future work can study the impact that B_φ plays regarding the driving mechanism of these wide angle outflows. Lastly, in chapter 4 the OWA-ESN was implemented in forecasting the rate of mass accretion in the late phases of protostellar evolution. The implementation involved splitting the simulation data into 3 segments to initialize the OWA-ESN and output predictions on future values. Future work will aim to train and validate the network on multiple sets of data and be used to predict the approximate time of disk formation as discussed in subsection 21.3.

BIBLIOGRAPHY

- [1] Michaël Gillon.
A novel seti strategy targeting the solar focal regions of the most nearby stars.
Acta Astronautica, 94(2):629–633, 2014.
- [2] Mark R Krumholz, Blakesley Burkhart, John C Forbes, and Roland M Crocker.
A unified model for galactic discs: star formation, turbulence driving, and mass transport.
Monthly Notices of the Royal Astronomical Society, 477(2):2716–2740, 2018.
- [3] Frank H Shu, Fred C Adams, and Susana Lizano.
Star formation in molecular clouds: observation and theory.
Annual review of astronomy and astrophysics, 25(1):23–81, 1987.
- [4] Ryo Kandori, Motohide Tamura, Nobuhiko Kusakabe, Yasushi Nakajima, Jungmi Kwon, Takahiro Nagayama, Tetsuya Nagata, Kohji Tomisaka, and Ken, Aïichi Tatematsu.
Distortion of magnetic fields in a starless core: Near-infrared polarimetry of fest 1–457.
The Astrophysical Journal, 845(1):32, 2017.
- [5] Shantanu Basu, Glenn E Ciolek, and James Wurster.
Nonlinear evolution of gravitational fragmentation regulated by magnetic fields and ambipolar diffusion.
New Astronomy, 14(3):221–237, 2009.
- [6] Josep M Girart, Ramprasad Rao, and Daniel P Marrone.
Magnetic fields in the formation of sun-like stars.
Science, 313(5788):812–814, 2006.
- [7] Josep M Girart, Maria T Beltrán, Qizhou Zhang, Ramprasad Rao, and Robert Estalella.
Magnetic fields in the formation of massive stars.
science, 324(5933):1408–1411, 2009.
- [8] Leverett Davis Jr and Jesse L Greenstein.
The polarization of starlight by aligned dust grains.
The Astrophysical Journal, 114:206, 1951.

BIBLIOGRAPHY

- [9] Bartek Ewertowski and Shantanu Basu.
A mathematical model for an hourglass magnetic field.
The Astrophysical Journal, 767(1):33, 2013.
- [10] HC Spruit.
Essential magnetohydrodynamics for astrophysics.
arXiv preprint arXiv:1301.5572, 2013.
- [11] Richard M Crutcher.
Magnetic fields in molecular clouds.
Annual Review of Astronomy and Astrophysics, 50:29–63, 2012.
- [12] Yuri Aikawa, Valentine Wakelam, Robin T Garrod, and Eric Herbst.
Molecular evolution and star formation: from prestellar cores to protostellar cores.
The Astrophysical Journal, 674(2):984, 2008.
- [13] Edwin A Bergin and Mario Tafalla.
Cold dark clouds: the initial conditions for star formation.
Annu. Rev. Astron. Astrophys., 45:339–396, 2007.
- [14] EI Vorobyov and Shantanu Basu.
The origin of episodic accretion bursts in the early stages of star formation.
The Astrophysical Journal Letters, 633(2):L137, 2005.
- [15] Gianfranco Bino and Shantanu Basu.
Fitting an analytic magnetic field to a prestellar core.
The Astrophysical Journal, 911(1):15, 2021.
- [16] L Mestel and L Spitzer Jr.
Star formation in magnetic dust clouds.
Monthly Notices of the Royal Astronomical Society, 116(5):503–514, 1956.
- [17] Sayantan Auddy, Shantanu Basu, and Takahiro Kudoh.
The transition from a lognormal to a power-law column density distribution in molecular clouds: An imprint of the initial magnetic field and turbulence.
The Astrophysical Journal Letters, 881(1):L15, 2019.
- [18] Philip C Myers, Shantanu Basu, and Sayantan Auddy.
Magnetic field structure in spheroidal star-forming clouds.
The Astrophysical Journal, 868(1):51, 2018.
- [19] Philip C Myers, Ian W Stephens, Sayantan Auddy, Shantanu Basu, Tyler L Bourke, and Charles LH Hull.

- Magnetic field structure in spheroidal star-forming clouds. ii. estimating field structure from observed maps.
The Astrophysical Journal, 896(2):163, 2020.
- [20] Subrahmanyan Chandrasekhar and Enrico Fermi.
Magnetic fields in spiral arms.
The Astrophysical Journal, 118:113, 1953.
- [21] Ryo Kandori, Yasushi Nakajima, Motohide Tamura, Ken, Aïichi Tatematsu, Yuri Aikawa, Takahiro Naoi, Koji Sugitani, Hidehiko Nakaya, Takahiro Nagayama, Tetsuya Nagata, et al.
Near-infrared imaging survey of bok globules: Density structure.
The Astronomical Journal, 130(5):2166, 2005.
- [22] Ryo Kandori, Kohji Tomisaka, Motohide Tamura, Masao Saito, Nobuhiko Kusakabe, Yasushi Nakajima, Jungmi Kwon, Takahiro Nagayama, Tetsuya Nagata, and Ken, Aïichi Tatematsu.
Distortion of magnetic fields in a starless core. iv. magnetic field scaling on density and mass-to-flux ratio distribution in fest 1-457.
The Astrophysical Journal, 865(2):121, 2018.
- [23] L Spitzer Jr and TA McGlynn.
Disorientation of interstellar grains in suprathemal rotation.
The Astrophysical Journal, 231:417–424, 1979.
- [24] A Lazarian and WG Roberge.
Barnett relaxation in thermally rotating grains.
The Astrophysical Journal, 484(1):230, 1997.
- [25] A Lazarian.
Gold-type mechanisms of grain alignment.
Monthly Notices of the Royal Astronomical Society, 268(3):713–723, 1994.
- [26] Thiem Hoang and Alex Lazarian.
Radiative torque alignment: essential physical processes.
Monthly Notices of the Royal Astronomical Society, 388(1):117–143, 2008.
- [27] PG Martin.
Interstellar polarization from a medium with changing grain alignment.
The Astrophysical Journal, 187:461–472, 1974.
- [28] BG Andersson, A Lazarian, and John E Vaillancourt.
Interstellar dust grain alignment.

- Annual Review of Astronomy and Astrophysics*, 53:501–539, 2015.
- [29] Melissa L Enoch, Jeong-Eun Lee, Paul Harvey, Michael M Dunham, and Scott Schnee.
A candidate detection of the first hydrostatic core.
The Astrophysical Journal Letters, 722(1):L33, 2010.
- [30] KL Thompson, TH Troland, and C Heiles.
A survey of magnetic field strengths in the envelopes of molecular clouds via the 18 cm oh
zeeman effect.
The Astrophysical Journal, 884(1):49, 2019.
- [31] Katia Ferrière.
Interstellar magnetic fields: from galactic scales to the edge of the heliosphere.
577:012008, 2015.
- [32] Zhi-Yun Li and Fumitaka Nakamura.
Magnetically regulated star formation in turbulent clouds.
The Astrophysical Journal Letters, 609(2):L83, 2004.
- [33] Fumitaka Nakamura and Zhi-Yun Li.
Quiescent cores and the efficiency of turbulence-accelerated, magnetically regulated star
formation.
The Astrophysical Journal, 631(1):411, 2005.
- [34] Takahiro Kudoh and Shantanu Basu.
Three-dimensional simulation of magnetized cloud fragmentation induced by nonlinear
flows and ambipolar diffusion.
The Astrophysical Journal Letters, 679(2):L97, 2008.
- [35] Shantanu Basu, Glenn E Ciolek, Wolf B Dapp, and James Wurster.
Magnetically-regulated fragmentation induced by nonlinear flows and ambipolar diffusion.
New Astronomy, 14(5):483–495, 2009.
- [36] Takahiro Kudoh and Shantanu Basu.
Induced core formation time in subcritical magnetic clouds by large-scale trans-alfvénic
flows.
The Astrophysical Journal, 794(2):127, 2014.
- [37] Sayantan Auddy, Shantanu Basu, and Takahiro Kudoh.
The effect of magnetic fields and ambipolar diffusion on the column density probability
distribution function in molecular clouds.
Monthly Notices of the Royal Astronomical Society, 474(1):400–410, 2018.

- [38] Telemachos Ch Mouschovias and Glenn E Ciolek.
Magnetic fields and star formation: a theory reaching adulthood.
pages 305–340, 1999.
- [39] Mark R Krumholz and Christoph Federrath.
The role of magnetic fields in setting the star formation rate and the initial mass function.
Frontiers in Astronomy and Space Sciences, 6:7, 2019.
- [40] Nigel R Minchin, Vitor HR Bonifacio, and Alexander G Murray.
Submillimetre polarimetric observations of s140 and gl 2591: investigating the role of
viewing angle on observed polarization position angles.
Astronomy and Astrophysics, 315:L5–L8, 1996.
- [41] Jason Glenn, Christopher K Walker, and Erick T Young.
Magnetic fields in star formation regions: 1.3 millimeter continuum polarimetry.
The Astrophysical Journal, 511(2):812, 1999.
- [42] Jacques P Vallée, Pierre Bastien, and Jane S Greaves.
Highly polarized thermal dust emission in the bok globule cb 068.
The Astrophysical Journal, 542(1):352, 2000.
- [43] Tomoaki Matsumoto, Masahiro N Machida, Kohji Tomisaka, and Tomoyuki Hanawa.
Self-gravitational collapse of a magnetized cloud core: high resolution simulations with
three-dimensional mhd nested grid.
Computer physics communications, 164(1-3):229–236, 2004.
- [44] Masahiro N Machida, Tomoaki Matsumoto, Kohji Tomisaka, and Tomoyuki Hanawa.
Collapse and fragmentation of rotating magnetized clouds,Äii. magnetic flux-spin relation.
Monthly Notices of the Royal Astronomical Society, 362(2):369–381, 2005.
- [45] Masahiro N Machida, Tomoaki Matsumoto, Tomoyuki Hanawa, and Kohji Tomisaka.
Collapse and fragmentation of rotating magnetized clouds,Äiii. binary formation and
fragmentation of first cores.
Monthly Notices of the Royal Astronomical Society, 362(2):382–402, 2005.
- [46] Masahiro N Machida, Shu-ichiro Inutsuka, and Tomoaki Matsumoto.
Second core formation and high-speed jets: resistive magnetohydrodynamic nested grid
simulations.
The Astrophysical Journal Letters, 647(2):L151, 2006.
- [47] Masahiro N Machida, Shu-ichiro Inutsuka, and Tomoaki Matsumoto.
Magnetic fields and rotations of protostars.
The Astrophysical Journal, 670(2):1198, 2007.

- [48] Takenori Nakano, Ryoichi Nishi, and Toyoharu Umebayashi.
Mechanism of magnetic flux loss in molecular clouds.
The Astrophysical Journal, 573(1):199, 2002.
- [49] Masahiro N Machida, Shu-ichiro Inutsuka, and Tomoaki Matsumoto.
First direct simulation of brown dwarf formation in a compact cloud core.
The Astrophysical Journal Letters, 699(2):L157, 2009.
- [50] Naoya Fukuda and Tomoyuki Hanawa.
Gravitational and parametric instabilities of the interstellar medium in which the alfvén wave travels.
The Astrophysical Journal, 517(1):226, 1999.
- [51] Tomoya Hirota, Richard L Plambeck, Melvyn CH Wright, Masahiro N Machida, Yuko Matsushita, Kazuhito Motogi, Mi Kyoung Kim, Ross A Burns, and Mareki Honma.
Magnetic field structure of orion source i.
arXiv preprint arXiv:2005.13077, 2020.
- [52] Peter Goldreich, Douglas A Keeley, and John Y Kwan.
Astrophysical masers. 11. polarization properties.
The Astrophysical Journal, 179:111–134, 1973.
- [53] Mantas Lukoševičius and Herbert Jaeger.
Reservoir computing approaches to recurrent neural network training.
Computer Science Review, 3(3):127–149, 2009.
- [54] Xuanlin Liu, Mingzhe Chen, Changchuan Yin, and Walid Saad.
Analysis of memory capacity for deep echo state networks.
pages 443–448, 2018.
- [55] Qiuyi Wu, Ernest Fokoue, and Dhireesha Kudithipudi.
On the statistical challenges of echo state networks and some potential remedies.
arXiv preprint arXiv:1802.07369, 2018.
- [56] Eduard I Vorobyov, Vardan Elbakyan, Takashi Hosokawa, Yuya Sakurai, Manuel Guedel, and Harold Yorke.
Effect of accretion on the pre-main-sequence evolution of low-mass stars and brown dwarfs.
Astronomy & Astrophysics, 605:A77, 2017.
- [57] Eduard I Vorobyov and Shantanu Basu.
The burst mode of accretion and disk fragmentation in the early embedded stages of star formation.
The Astrophysical Journal, 719(2):1896, 2010.

- [58] Paola Caselli, Priscilla J Benson, Philip C Myers, and Mario Tafalla.
Dense cores in dark clouds. xiv. n2h+ (1-0) maps of dense cloud cores.
The Astrophysical Journal, 572(1):238, 2002.
- [59] D. Ruelle J. P. Eckmann, S. Oliffson Kamphorst and S. Ciliberto.
Liapunov exponents from time series.
Phys. Rev. A 34, 4971, 1986.
- [60] Jaideep Pathak, Brian Hunt, Michelle Girvan, Zhixin Lu, and Edward Ott.
Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir
computing approach.
Physical review letters, 120(2):024102, 2018.
- [61] Claudio Gallicchio and Alessio Micheli.
Deep echo state network (deepesn): A brief survey.
arXiv preprint arXiv:1712.04323, 2017.
- [62] Jingpei Dan, Wenbo Guo, Weiren Shi, Bin Fang, and Tingping Zhang.
Deterministic echo state networks based stock price forecasting.
2014, 2014.
- [63] Xiao-chuan Sun, Hong-yan Cui, Ren-ping Liu, Jian-ya Chen, and Yun-jie Liu.
Modeling deterministic echo state network with loop reservoir.
Journal of Zhejiang University SCIENCE C, 13(9):689–701, 2012.
- [64] Eduard I. Vorobyov and Shantanu Basu.
Variable protostellar accretion with episodic bursts.
The Astrophysical Journal, 805(2):115, may 2015.
doi: 10.1088/0004-637x/805/2/115.
URL <https://doi.org/10.1088/0004-637x/805/2/115>.
- [65] Marc Audard, Péter Abrahám, Michael M Dunham, Joel D Green, Nicolas Grosso, Kenji
Hamaguchi, Joel H Kastner, Agnes Kóspál, Giuseppe Lodato, Marina M Romanova,
et al.
Episodic accretion in young stars.
Protostars and Planets VI, 387(1401.3368), 2014.
- [66] Luca Anthony Thiede and Ulrich Parlitz.
Gradient based hyperparameter optimization in echo state networks.
Neural Networks, 115:23–29, 2019.
- [67] Erich Zauderer.
Partial differential equations of applied mathematics, volume 71.

- John Wiley & Sons, 2011.
- [68] Paul T Boggs and Jon W Tolle.
Sequential quadratic programming.
Acta numerica, 4(1):1–51, 1995.
- [69] W. Murray M.A. Saunders Gill, P.E. and M.H. Wright.
Procedures for Optimization Problems with a Mixture of Bounds and General Linear Constraints, volume 10.
ACM Trans. Math. Software, 1984.
- [70] Constrained nonlinear optimization algorithms.
URL <https://www.mathworks.com/help/optim/>.
- [71] Tomoaki Matsumoto and Tomoyuki Hanawa.
A fast algorithm for solving the poisson equation on a nested grid.
The Astrophysical Journal, 583(1):296, 2003.
- [72] Saul A Teukolsky, Brian P Flannery, WH Press, and WT Vetterling.
Numerical recipes in c.
SMR, 693(1):59–70, 1992.

Appendix

Chapter II

A1.1 Magnetic Field Model Mathematical Derivation

The total magnetic field is assumed to take the following form:

$$(A1.1) \quad \mathbf{B} = \mathbf{B}_c + \mathbf{B}_0$$

for a background field \mathbf{B}_0 assumed to be taken along the z -direction:

$$(A1.2) \quad \mathbf{B}_0 \equiv B_0 \hat{\mathbf{z}}$$

and \mathbf{B}_c is the local magnetic field generated within the core. Since the model is assumed to be axisymmetric, the magnetic field $B \equiv B(r, z)$ becomes independent on the azimuthal angle φ and leads to a purely poloidal structure:

$$(A1.3) \quad \mathbf{B} = B_r \hat{\mathbf{r}} + B_z \hat{\mathbf{z}}$$

The magnetic vector potential \mathbf{A} is related to the magnetic field as:

$$(A1.4) \quad \mathbf{B} = \nabla \times \mathbf{A}$$

where the cross product suggests that the magnetic vector potential is purely toroidal and given as:

$$(A1.5) \quad \mathbf{A} = A \hat{\boldsymbol{\varphi}}$$

The magnetic vector potential is related to the magnetic flux given through the surface integral given be:

$$(A1.6) \quad \Phi(r, z) = \int_S \mathbf{B} \cdot d\mathbf{S} = 2\pi \int_0^r \mathbf{B} \cdot \hat{\mathbf{z}} r' dr'$$

Define $\tilde{\Phi} \equiv \Phi/2\pi$ and insert equation A1.4:

$$\begin{aligned} \tilde{\Phi}(r, z) &= \int_0^r (\nabla \times \mathbf{A}) \cdot \hat{\mathbf{z}} r' dr' \\ &= \int_0^r \left(\frac{1}{r'} \frac{\partial A_z}{\partial \varphi} - \frac{\partial A_\varphi}{\partial z} \right) \hat{\mathbf{r}}' + \left(\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r'} \right) \hat{\boldsymbol{\varphi}} + \frac{1}{r'} \left(\frac{\partial}{\partial r'} (r' A_\varphi) - \frac{\partial A_r}{\partial \varphi} \right) \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} r' dr' \\ &= \int_0^r \left(\frac{\partial}{\partial r'} (r' A_\varphi) \right) dr' \\ &= r A_\varphi \end{aligned}$$

and so magnetic vector potential is taken as:

$$(A1.7) \quad \mathbf{A} = \frac{\tilde{\Phi}}{r} \hat{\boldsymbol{\varphi}}$$

The Ampere-Maxwell law gives the relation of:

$$\begin{aligned}\frac{4\pi}{c}\mathbf{j} + \nabla^2\mathbf{A} &= 0 \\ \nabla \cdot \mathbf{A} &= 0\end{aligned}$$

where \mathbf{j} is a current density. Since the magnetic vector potential is along the φ -direction, it follows that the current density also runs in the φ -direction and so:

$$(A1.8) \quad \frac{4\pi}{c}j + \nabla^2 A = \frac{A}{r^2}$$

Define the operator $\tilde{\nabla}$ as:

$$\tilde{\nabla} = \frac{\partial}{\partial r}\hat{\mathbf{r}} + \frac{\partial}{\partial z}\hat{\mathbf{z}}$$

and express equation A1.8 in terms of $\tilde{\nabla}$:

$$(A1.9) \quad \begin{aligned}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial A}{\partial r}\right)\hat{\mathbf{r}} + \frac{\partial^2 A}{\partial z^2}\hat{\mathbf{z}} &= \frac{A}{r^2} - \frac{4\pi}{c}j \\ \tilde{\nabla} \cdot \left(r\frac{\partial A}{\partial r}\hat{\mathbf{r}} + r\frac{\partial A}{\partial z}\hat{\mathbf{z}}\right) &= \frac{A}{r} - \frac{4\pi r}{c}j \\ -\tilde{\nabla} \cdot (r\tilde{\nabla}A) + \frac{A}{r} &= \frac{4\pi r}{c}j\end{aligned}$$

where the left hand side is expressed as a self adjoint elliptic operator L :

$$(A1.10) \quad LA = -\tilde{\nabla} \cdot (r\tilde{\nabla}A) + \frac{A}{r} = \frac{4\pi r}{c}j$$

The model imposes the conditions that the magnetic field calms to B_0 outside the core and thus one expects no flux from the local field outside the core

$$\tilde{\Phi}(0, z) = \tilde{\Phi}(R, z) = \lim_{z \rightarrow \pm\infty} \tilde{\Phi}(r, z) = 0$$

using the relation $A = \tilde{\Phi}/r$ gives the vector potential counterpart to these boundary conditions as:

$$\begin{aligned}A(0, z) &= \text{bounded} \\ A(R, z) &= 0 \\ \lim_{z \rightarrow \pm\infty} A(r, z) &= 0\end{aligned}$$

Equation A1.9 is solved using a Green's function that satisfies:

$$(A1.11) \quad LG(\xi, \eta, r, z) = \delta(r - \xi)\delta(z - \eta)$$

for a point source located at $(r, z) = (\xi, \eta) \in D$ where $D = \{(r, z) \in \mathbb{R}^2 | r \in (0, R), z \in (-\infty, \infty)\}$. Equation A1.11 can be written as:

$$LG = -\tilde{\nabla} \cdot (r\tilde{\nabla}G) + \frac{G}{r} = \delta(r - \xi)\delta(z - \eta)$$

$$(A1.12) \quad LG = -r \frac{\partial^2 G}{\partial z^2} + L_r G = \delta(r - \xi) \delta(z - \eta)$$

where the operator L_r has been defined as:

$$(A1.13) \quad L_r G = -\frac{\partial}{\partial r} \left(r \frac{\partial G}{\partial r} \right) + \frac{G}{r}$$

The Green's function is expanded in a series of eigenfunctions as:

$$(A1.14) \quad G(\xi, \eta, r, z) = \sum_{m=1}^{\infty} C_m(\xi, \eta, z) \psi_m(r)$$

and the operator L_r is represented as a Sturm-Liouville equation for eigenfunctions $\psi_m(r)$:

$$(A1.15) \quad -\frac{\partial}{\partial r} \left[r \frac{d\psi_m(r)}{dr} \right] + \frac{1}{r} \psi_m(r) = \lambda w(r) \psi_m(r)$$

The Sturm-Liouville form of Bessel's equation is given as

$$-\frac{d}{dx} \left[x \frac{dy(x)}{dx} \right] + \frac{n^2}{x} y(x) = \lambda x y(x)$$

which has the equivalent form:

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2) y = 0$$

Define the scaled variable $v = x/k$ and substitute:

$$\begin{aligned} (kv)^2 \frac{1}{k^2} \frac{d^2 y}{dv^2} + kv \frac{1}{k} \frac{dy}{dv} + (k^2 v^2 - n^2) y &= 0 \\ v \frac{d^2 y}{dv^2} + \frac{dy}{dv} + \left(k^2 v - \frac{n^2}{v} \right) y &= 0 \\ -\frac{d}{dv} \left(v \frac{dy}{dv} \right) + y \frac{n^2}{v} &= y k^2 v \end{aligned}$$

By comparison, equation A1.15 can be taken as a Bessel Sturm-Liouville equation for the scaled variable $v = x/k$ [67]. Thus for equation A1.15, $w(r) = r$, $n = 1$ and $k^2 = \lambda_m$ to give:

$$(A1.16) \quad -\frac{\partial}{\partial r} \left[r \frac{d\psi_m(r)}{dr} \right] + \frac{1}{r} \psi_m(r) = \lambda_m r \psi_m(r)$$

$$(A1.17) \quad L_r \psi_m = \lambda_m r \psi_m$$

which have the boundary conditions for $0 \leq r \leq R$:

$$\begin{aligned} \psi_m(0) &= \text{bounded} \\ \psi_m(R) &= 0 \end{aligned}$$

The eigenvalues are thus found by seeking a solution to the scaled Bessel differential equation which has a general solution of the form

$$(A1.18) \quad \psi_m(r) = C_1 J_n(kr) + C_2 Y_n(kr)$$

Applying the boundary condition $\psi_m(R) = 0$

$$\psi_m(R) = C_1 J_n(kR) + C_2 Y_n(kR) = 0$$

the second boundary condition requires $\psi_m(0)$ to be bounded and since a singular value comes from $Y_n(0)$, it's required that $C_2 = 0$ and so in order to satisfy $\psi_m(R) = 0$ it is required that

$$J_n(kR) = 0$$

recalling that equation A1.16 had $n = 1$ and $k^2 = \lambda_m$ the eigenvalue of the Sturm-Liouville equation is given as:

$$(A1.19) \quad \lambda_m = \left(\frac{a_{m,1}}{R} \right)^2$$

where $a_{m,1}$ is the m^{th} root to the Bessel function of first kind. Since $J_n(kr)$ is a solution to a Sturm-Liouville equation, it follows that the functions are orthogonal:

$$(\psi_n, \psi_m) = \int_0^R \psi_n(r) \psi_m r dr = \delta_{nm}$$

and an orthonormal set of eigenfunctions are given as:

$$\psi_m(r) = \frac{\hat{\psi}_m(r)}{\|\psi_m(r)\|} = \frac{J_n(kr)}{\|J_n(kr)\|}$$

where $\|J_n(kr)\|^2$ is given to as [67]:

$$\|J_n(kr)\|^2 = \int_0^R r J_n^2(kr) dr = \frac{R^2}{2} J_{n+1}^2(kR)$$

Therefore, the eigenfunctions of the Sturm-Liouville equation are given as:

$$\psi_m(r) = \frac{\sqrt{2} J_n(kr)}{R |J_{n+1}(kR)|}$$

recalling that $n = 1$ and $k^2 = \lambda_m$ gives

$$(A1.20) \quad \psi_m(r) = \frac{\sqrt{2} J_1(\sqrt{\lambda_m} r)}{R |J_2(\sqrt{\lambda_m} R)|}$$

Making use of the orthonormal properties of the eigenfunctions, equation A1.11 can be written as an ordinary differential equation in terms of the expansion coefficients of the Green's function.

Multiplying both sides of equation A1.11 by $\psi_n(r)$ and integrating with respect to r from 0 to R :

$$\begin{aligned}
& \int_0^R \psi_n(r) \cdot \left(-r \frac{\partial^2 G}{\partial z^2} + L_r G \right) dr = \int_0^R \psi_n(r) \cdot \delta(r - \xi) \delta(z - \eta) dr \\
& \int_0^R \psi_n(r) \cdot \left(-r \frac{\partial^2}{\partial z^2} C_m(\xi, \eta, z) \psi_m(r) + L_r C_m(\xi, \eta, z) \psi_m(r) \right) dr = \int_0^R \psi_n(r) \cdot \delta(r - \xi) \delta(z - \eta) dr \\
& - \int_0^R \psi_n(r) r \frac{\partial^2}{\partial z^2} C_m(\xi, \eta, z) \psi_m(r) dr + \int_0^R \psi_n(r) L_r C_m(\xi, \eta, z) \psi_m(r) dr = \int_0^R \psi_n(r) \cdot \delta(r - \xi) \delta(z - \eta) dr \\
& - \frac{\partial^2 C_m(\xi, \eta, z)}{\partial z^2} \int_0^R \psi_n(r) \psi_m(r) r dr + C_m(\xi, \eta, z) \int_0^R \psi_n(r) L_r \psi_m(r) dr = \delta(z - \eta) \int_0^R \psi_n(r) \delta(r - \xi) dr \\
& - \frac{\partial^2 C_m(\xi, \eta, z)}{\partial z^2} + C_m(\xi, \eta, z) \int_0^R \psi_n(r) L_r \psi_m(r) dr = \delta(z - \eta) \int_0^R \psi_n(r) \delta(r - \xi) dr
\end{aligned}$$

The right hand side is simplified by one of the properties of the Dirac delta function:

$$\int_{\xi-\epsilon}^{\xi+\epsilon} f(r) \delta(r - \xi) dr = f(\xi)$$

and so:

$$\text{(A1.21)} \quad - \frac{\partial^2 C_m(\xi, \eta, z)}{\partial z^2} + C_m(\xi, \eta, z) \int_0^R \psi_n(r) L_r \psi_m(r) dr = \delta(z - \eta) \psi_n(\xi)$$

The remaining integral can be found by using the eigenvalue and eigenfunction properties of the Sturm-Liouville equation [67] that states if one has two distinct eigenvalues λ_i and λ_j corresponding to two eigenfunctions v_i and v_j then:

$$\int_0^L [v_i L v_j - v_j L v_i] dx = (\lambda_i - \lambda_j) \int_0^L w v_i v_j dx$$

translating this to be in the form of the integral in equation A1.21 gives:

$$\int_0^R \psi_n(r) L_r \psi_m(r) dr = \lambda_n \int_0^R w(r) \psi_n(r) \psi_m(r) dr$$

and so:

$$- \frac{\partial^2 C_m(\xi, \eta, z)}{\partial z^2} + C_m(\xi, \eta, z) \lambda_n \int_0^R w(r) \psi_n(r) \psi_m(r) dr = \delta(z - \eta) \psi_n(\xi)$$

where $w(r) = r$ so:

$$\begin{aligned}
& - \frac{\partial^2 C_m(\xi, \eta, z)}{\partial z^2} + C_m(\xi, \eta, z) \lambda_n \int_0^R \psi_n(r) \psi_m(r) r dr = \delta(z - \eta) \psi_n(\xi) \\
\text{(A1.22)} \quad & - \frac{\partial^2 C_n(\xi, \eta, z)}{\partial z^2} + C_n(\xi, \eta, z) \lambda_n = \delta(z - \eta) \psi_n(\xi)
\end{aligned}$$

Equation A1.22 gives ψ_n as a function of ξ and so it follows from this that the expansion coefficients are separable:

$$C_n(\xi, \eta, z) = K_n(\eta, z) \psi_n(\xi)$$

which is substituted back into equation A1.22 to give:

$$(A1.23) \quad \begin{aligned} & -\psi_n(\xi) \frac{\partial^2 K_n(\eta, z)}{\partial z^2} + \psi_n(\xi) K_n(\eta, z) \lambda_n = \delta(z - \eta) \psi_n(\xi) \\ & -\frac{\partial^2 K_n(\eta, z)}{\partial z^2} + K_n(\eta, z) \lambda_n = \delta(z - \eta) \end{aligned}$$

Equation A1.23 is solved by using the Fourier transform method [67] for differential equations of the form

$$y''(x) - k^2 y(x) = -f(x)$$

Take the Fourier transform of some function $y(x)$ to be $Y(\omega)$ given as:

$$Y(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega x} y(x) dx$$

Thus the Fourier transform of $K_n(\eta, z)$ is given as:

$$(A1.24) \quad \kappa_n(\eta, \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega z} K_n(\eta, z) dz$$

A solution is found by multiplying equation A1.23 by $(1/\sqrt{2\pi})e^{i\omega z}$ and integrating with respect to z from $-\infty$ to $+\infty$:

$$(A1.25) \quad \begin{aligned} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{i\omega z} \delta(z - \eta) dz &= -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{i\omega z} \frac{\partial^2 K_n(\eta, z)}{\partial z^2} dz + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{i\omega z} K_n(\eta, z) \lambda_n dz \\ &= -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega z} \frac{\partial^2 K_n(\eta, z)}{\partial z^2} dz + \lambda_n \kappa(\eta, \omega) \end{aligned}$$

Using the Fourier transform property [67]:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega x} f^{(n)}(x) dx = -(i\omega)^n \mathcal{F}(\omega)$$

where $\mathcal{F}(\omega)$ is the Fourier transform to $f(x)$, equation A1.25 becomes:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega z} \delta(z - \eta) dz = -(i\omega)^2 \kappa(\eta, \omega) + \lambda_n \kappa(\eta, \omega)$$

Define the Fourier transform of the left hand side to be $D(\eta, \omega)$ and so:

$$D(\eta, \omega) = -(i\omega)^2 \kappa(\eta, \omega) + \lambda_n \kappa(\eta, \omega) = (\omega^2 + \lambda_n) \kappa(\eta, \omega)$$

where solving for $\kappa(\eta, \omega)$ gives:

$$\kappa(\eta, \omega) = \frac{D(\eta, \omega)}{(\omega^2 + \lambda_n)}$$

Define $F(\omega) = 1/(\omega^2 + \lambda_n)$ and so:

$$(A1.26) \quad \kappa(\eta, \omega) = D(\eta, \omega) F(\omega)$$

where the solution can be found by taking the inverse transform. In order to do this it is required that the inverse transform of the functions $D(\eta, \omega)$ and $F(\omega)$ are known. The inverse of $D(\eta, \omega)$ is just $y(z) = \delta(z - \eta)$ and the inverse transform of $F(\omega)$ is:

$$f(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega z} F(\omega) d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-i\omega z}}{\omega^2 + \lambda_n} d\omega$$

Using Mathematica's symbolic integrator, it is found that $f(z)$ is:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-i\omega z}}{\omega^2 + \lambda_n} d\omega = \frac{\sqrt{2\pi}}{2\sqrt{\lambda_n}} e^{-\sqrt{\lambda_n}|z|}$$

and so by convolution theorem [67], the inverse transform of $\kappa(\eta, \omega)$ is:

$$K_n(\eta, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} y(t) \int_{-\infty}^{\infty} e^{-i\omega(z-t)} F(\omega) d\omega dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(t) f(z-t) dt$$

substituting in:

$$f(z-t) = \frac{\sqrt{2\pi}}{2\sqrt{\lambda_n}} e^{-\sqrt{\lambda_n}|z-t|}$$

gives:

$$K_n(\eta, z) = \frac{1}{2\sqrt{\lambda_n}} \int_{-\infty}^{\infty} \delta(t-\eta) e^{-\sqrt{\lambda_n}|z-t|} dt$$

which can be evaluated using the same Dirac delta property of equation A1.21:

$$(A1.27) \quad K_n(\eta, z) = \frac{1}{2\sqrt{\lambda_n}} e^{-\sqrt{\lambda_n}|z-\eta|}$$

Therefore, the Green's function can be written as:

$$(A1.28) \quad \begin{aligned} G(\xi, \eta, r, z) &= \sum_{m=1}^{\infty} K_m(\eta, z) \psi_m(\xi) \frac{\sqrt{2} J_1(\sqrt{\lambda_m} r)}{R |J_2(\sqrt{\lambda_m} R)|} \\ &= \sum_{m=1}^{\infty} K_m(\eta, z) \frac{2 J_1(\sqrt{\lambda_m} \xi) J_1(\sqrt{\lambda_m} r)}{R^2 |J_2(\sqrt{\lambda_m} R)|^2} \\ &= \sum_{m=1}^{\infty} \frac{J_1(\sqrt{\lambda_m} \xi) J_1(\sqrt{\lambda_m} r) e^{-\sqrt{\lambda_m}|z-\eta|}}{R^2 \sqrt{\lambda_m} |J_2(\sqrt{\lambda_m} R)|^2} \end{aligned}$$

which satisfies the general set of boundary conditions:

$$\alpha G(\xi, \eta, r, z) + \beta \frac{\partial G}{\partial r} \Big|_{\partial G} = B$$

Where the functions are all evaluated at the boundaries. Defining the case where $\alpha \neq 0$ and $\beta = 0$ as a boundary condition of the first kind, $\alpha = 0$ and $\beta \neq 0$ as a boundary condition of the second kind and lastly $\alpha \neq 0$ and $\beta \neq 0$ as a boundary condition of the third kind. One can also have mixed boundary conditions where the conditions are combinations of the three. The domain is generally split into three subsets S_1 , S_2 and S_3 satisfying boundary conditions of the

first, second and third kind respectively. The solution to an elliptic differential equation given the Green's function takes the form [67]:

$$u(r, z) = \int \int_G \rho(\xi) G(\xi, \eta, r, z) F(\xi, \eta) d\Omega - \int_{S_1} \frac{p(r)B}{\alpha} \frac{\partial G}{\partial r} dS + \int_{S_2 \cup S_3} \frac{p(r)B}{\beta} G(\xi, \eta, r, z) dS$$

The conditions set earlier are Dirichlet boundary conditions and so it follows that they are of the first kind so the regions S_2 and S_3 can be ignored:

$$u(r, z) = \int \int_G \rho(\xi) G(\xi, \eta, r, z) F(\xi, \eta) d\Omega - \int_{S_1} \frac{p(r)B}{\alpha} \frac{\partial G}{\partial r} dS$$

and so $\alpha = 1$ and $B = 0$ for this model's boundary conditions

$$(A1.29) \quad u(r, z) = \int \int_G \rho(\xi) G(\xi, \eta, r, z) F(\xi, \eta) d\Omega$$

Equation A1.9 suggests that $\rho(\xi) = \xi$ and $F(\xi, \eta) = (4\pi/c)j(\xi, \eta)$ therefore the expression for $A(r, z)$ is given as:

$$(A1.30) \quad A(r, z) = \frac{4\pi}{c} \int_{-\infty}^{\infty} \int_0^R G(\xi, \eta, r, z) j(\xi, \eta) \xi d\xi d\eta$$

where it is assumed that the current density $j(r, z)$ is separable in the form:

$$j(r, z) = f(r)g(z)$$

where $g(z)$ is given as the current density along the vertical axis of a cylindrical coordinate system. This model selects $g(z)$ to be Gaussian of the form:

$$g(z) = e^{-z^2/h^2}$$

which is bell shaped and concentrates the bulk of the current along the equatorial plane $z = 0$. This allows for magnetic field lines that are smooth and do not exhibit sharp cusps as they would had one used a function like a Dirac delta. Here, h is a free parameter that is given to represent the model's scale height. Having expressions for the Green's function and the current density, equation A1.30 can be expressed as:

$$(A1.31) \quad A(r, z) = \frac{4\pi}{c} \sum_{m=1}^{\infty} \int_{-\infty}^{\infty} \int_0^R \frac{J_1(\sqrt{\lambda_m}\xi) J_1(\sqrt{\lambda_m}r) e^{-\sqrt{\lambda_m}|z-\eta|}}{R^2 \sqrt{\lambda_m} |J_2(\sqrt{\lambda_m}R)|^2} f(\xi) g(\eta) \xi d\xi d\eta$$

where the coefficients k_m are defined to be:

$$(A1.32) \quad k_m = \frac{2h\pi^{3/2} e^{h^2\lambda_m/4}}{cR^2 \sqrt{\lambda_m} [J_2(\sqrt{\lambda_m}R)]^2} \int_0^R f(\xi) J_1(\sqrt{\lambda_m}\xi) \xi d\xi$$

and so:

$$(A1.33) \quad A(r, z) = \frac{2}{h\sqrt{\pi}} \sum_{m=1}^{\infty} k_m e^{-h^2\lambda_m/4} J_1(\sqrt{\lambda_m}r) \int_{-\infty}^{\infty} g(\eta) e^{-\sqrt{\lambda_m}|z-\eta|} d\eta$$

The integral

$$\int_{-\infty}^{\infty} g(\eta) e^{-\sqrt{\lambda_m} |z-\eta|} d\eta$$

is evaluated by expressing it through a unilateral Laplace transform from the u domain to the $\sqrt{\lambda_m}$ domain:

$$\begin{aligned} \int_{-\infty}^{\infty} g(\eta) e^{-\sqrt{\lambda_m} |z-\eta|} d\eta &= \int_{-\infty}^0 g(\eta) e^{-\sqrt{\lambda_m} [-(z-\eta)]} d\eta + \int_0^{\infty} g(\eta) e^{-\sqrt{\lambda_m} (z-\eta)} d\eta \\ &= \mathcal{L} \{g(z-u)\} + \mathcal{L} \{g(z+u)\} \\ (A1.34) \quad &= \mathcal{L} \{g(z-u) + g(z+u)\} \end{aligned}$$

where

$$\mathcal{L} \{f(u)\} = \int_0^{\infty} f(u) e^{-u\sqrt{\lambda_m}} du$$

The integrals are evaluated in Mathematica to give:

$$\int_{-\infty}^{\infty} g(\eta) e^{-\sqrt{\lambda_m} |z-\eta|} d\eta = \frac{h\sqrt{\pi}}{2} e^{-h^2\lambda_m/4} e^{-z\sqrt{\lambda_m}} \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} - \frac{z}{h} \right) + e^{2z\sqrt{\lambda_m}} \operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} + \frac{z}{h} \right) \right]$$

and so inserting into equation A1.33 gives the magnetic vector potential to be:

$$(A1.35) \quad A(r, z) = \sum_{m=1}^{\infty} k_m J_1(\sqrt{\lambda_m}r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} + \operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} \right]$$

Using the relation $\mathbf{B} = \nabla \times \mathbf{A}$, the magnetic field is given by:

$$\begin{aligned} B_r &= -\frac{\partial A_\varphi}{\partial z} \\ B_z &= \frac{1}{r} \left(\frac{\partial}{\partial r} (rA_\varphi) \right) + B_0 \end{aligned}$$

which yield the explicit expressions:

$$\begin{aligned} B_r(r, z) &= \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_1(\sqrt{\lambda_m}r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} - \operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} \right] \\ B_z(r, z) &= \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_0(\sqrt{\lambda_m}r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} + \operatorname{erfc} \left(\frac{\sqrt{\lambda_m}h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} \right] + B_0 \end{aligned}$$

A1.2 Dimensional Analysis for magnetic field solution

Given the equations:

$$B_r(r, z) = \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_1(\sqrt{\lambda_m} r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} - \operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} \right]$$

$$B_z(r, z) = \sum_{m=1}^{\infty} k_m \sqrt{\lambda_m} J_0(\sqrt{\lambda_m} r) \left[\operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} + \frac{z}{h} \right) e^{z\sqrt{\lambda_m}} + \operatorname{erfc} \left(\frac{\sqrt{\lambda_m} h}{2} - \frac{z}{h} \right) e^{-z\sqrt{\lambda_m}} \right] + B_0$$

the dimensional contribution comes from B , $\sqrt{\lambda_m}$ and k_m . All bessel function and complementary error function arguments must be dimensionless and so checking the dimensions from B , $\sqrt{\lambda_m}$ and k_m will suffice.

$$\dim(\mathbf{B}) = \frac{\text{g}^{1/2}}{\text{cm}^{1/2}\text{s}}$$

$$\dim(\sqrt{\lambda_m}) = \frac{1}{\text{cm}}$$

and the dimension of k_m can be found by looking at the integral form:

$$k_m = \frac{2h\pi^{3/2}e^{h^2\lambda_m/4}}{cR^2\sqrt{\lambda_m} [J_2(\sqrt{\lambda_m}R)]^2} \int_0^R f(\xi) J_1(\sqrt{\lambda_m}\xi) \xi d\xi$$

The dimensional contribution outside the integral comes from h , R^2 , $\sqrt{\lambda_m}$ and the speed of light c where as the integrand contribution comes from $f(\xi)$ and ξ . Recalling that the function $f(\xi)$ comes from the current density

$$j(r, z) = f(r)g(z)$$

$f(r)$ will have units of current density

$$\dim(f(\xi)) = \frac{\text{current}}{\text{length}} = \frac{\text{cm}^{3/2}\text{g}^{1/2}\text{s}^{-2}}{\text{cm}}$$

and so the integrand will have dimensions of

$$\dim(f(\xi)) \times \dim(\xi) = \frac{\text{cm}^{3/2}\text{g}^{1/2}\text{s}^{-2}}{\text{cm}} \times \text{cm} = \text{cm}^{3/2}\text{g}^{1/2}\text{s}^{-2}$$

The portion outside the integral:

$$\frac{\text{cm}}{\text{cm} \cdot \text{s}^{-1}\text{cm}^2\text{cm}^{-1}} = \frac{\text{s}}{\text{cm}}$$

and so the dimensions of k_m are:

$$\dim(k_m) = \frac{\text{s}}{\text{cm}} \times \text{cm}^{3/2}\text{g}^{1/2}\text{s}^{-2} = \text{cm}^{1/2}\text{g}^{1/2}\text{s}^{-1}$$

and so with this:

$$\dim(k_m) \times \dim(\sqrt{\lambda_m}) = \text{cm}^{-1/2}\text{g}^{1/2}\text{s}^{-1} = \dim(\mathbf{B})$$

and so the dimensions all check out and with this, the definition of β_m is justified by the dimension check of $\dim(k_m) \times \dim(\sqrt{\lambda_m})$.

A1.3 SQP Algorithm

The optimization is done by minimizing the sum of squared residuals

$$(A1.36) \quad \min_{\beta, \eta} \|\Delta(r, z)\|^2 = \operatorname{argmin}_{\beta, \eta} \sum_i \left\| \frac{B_r(r_i, z_i)}{B_z(r_i, z_i)} - \tan \phi_i \right\|^2$$

This minimization is done using the sequential quadratic programming algorithm (SQP). The SQP algorithm is one of the most effective methods for nonlinear constrained optimization. It is an iterative method that works by solving quadratic subproblems and is generally used for systems having significant nonlinearities [68]. The SQP algorithm is generally applied to problems of the form:

$$(A1.37) \quad \begin{aligned} & \min_x f(x) \\ & \text{subject to } h(x) = 0 \\ & j(x) \leq 0. \end{aligned}$$

for $x \in \mathbb{R}^n$ and where $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function, $h(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the equality function and $j(x) : \mathbb{R}^n \rightarrow \mathbb{R}^p$ is the inequality function. The minimization is done by implementing Matlab's SQP algorithm within the `fmincon` function. A brief summary of the process is discussed [69, 70] and applied. The SQP algorithm is an iterative process where local quadratic subproblems are formed and solved to form a search direction for a line search procedure in order to iterate a sequence x_k for $k \in \mathbb{N}_0$. The quadratic subproblems should reflect the local properties of the non-linear optimization problem with respect to the local iterate x_k . The quadratic subproblem forms a quadratic approximation to the Lagrangian function

$$(A1.38) \quad \mathcal{L}(x, \zeta) = f(x) + \sum_{i=1}^m \zeta_i \cdot g_i(x)$$

where $g_i(x)$ is the i^{th} row of a vector $\mathbf{g}(x)$ containing the constraints $h(x)$ and $j(x)$. Taking the local quadratic approximation as well as linearizing the nonlinear constraints at the iterate x_k takes the minimization problem as:

$$(A1.39) \quad \begin{aligned} & \min_{d(x) \in \mathbb{R}^n} \nabla f(x_k)^T d(x) + \frac{1}{2} d(x)^T B_k d(x) \\ & \text{subject to } g_i(x_k) + \nabla g_i(x_k)^T d(x) = 0, \quad i = 1, \dots, m_{\text{eq}} \\ & g_i(x_k) + \nabla g_i(x_k)^T d(x) \leq 0, \quad i = m_{\text{eq}} + 1, \dots, m \end{aligned}$$

where $d(x) := x - x_k$ and $B_k := H \{f(x_k)\}$ is the approximation to the Hessian of the Lagrangian. The non-linear problem A1.39 is solved by iteratively forming solutions to quadratic subproblems of the form

$$(A1.40) \quad \begin{aligned} & \min_{d(x) \in \mathbb{R}^n} \Pi(d) = \nabla f(x_k)^T d(x) + \frac{1}{2} d(x)^T B_k d(x) \\ & \text{subject to } A_i d(x) = -\hat{c}_i, \quad i = 1, \dots, m_{\text{eq}} \\ & A_i d(x) \leq -\hat{c}_i, \quad i = m_{\text{eq}} + 1, \dots, m \end{aligned}$$

for \hat{c}_i being a vector of active constraints and the rows A_i for a matrix $\mathbf{A}_{m \times n}$ contain their respective gradients. Let the solution $d_k \equiv d(x)$ at the iterate k be written as:

$$(A1.41) \quad d_k = Z_k d_z$$

where the matrix Z_k is formed from the last $m - c$ columns of the QR decomposition of the matrix \mathbf{A}_k^T for $c < m$ being the total number of constraints. Finally, the vector d_z is found through the minimization of

$$(A1.42) \quad \Pi(d) = \nabla f(x_k)^T Z_k d_z + \frac{1}{2} d_z^T Z_k^T B_k Z_k d_z$$

Therefore one solves $\nabla \Pi(d) = 0$ for d_z acting as a new search direction for the algorithm. This forms the linear system

$$(A1.43) \quad Z_k^T B_k Z_k d_z = -Z_k^T \nabla f(x_k)$$

A step is taken in the direction of d_z as:

$$x_{k+1} = x_k + \gamma d_z$$

where the step length γ is taken either as a unit step if the constraints are not violated or given by:

$$(A1.44) \quad \gamma = \min_{i \in \{1, \dots, m\}} \left(\frac{-(A_i x_k + \hat{c}_i)}{A_i d_z} \right)$$

which is defined for constraints not in the active set and which is effectively the distance to the constraint boundaries. It is computed to produce "sufficient" decrease in a merit function taken to be [70]:

$$(A1.45) \quad \mathcal{M}(x) = f(x) + \sum_{i=1}^n r^{(i)} \cdot g_i(x) + \sum_{i=n+1}^m r^{(i)} \cdot \max(0, g_i(x))$$

with a penalty parameter

$$(A1.46) \quad r_{k+1}^{(i)} = \max \left(\zeta_i, \frac{\zeta_i + r_k^{(i)}}{2} \right) \quad i = 1, \dots, m$$

initialized to

$$(A1.47) \quad r^{(i)} = \frac{\|\nabla f(x)\|}{\|g_i(x)\|}$$

The optimality conditions for A1.40 can also be expressed as a system of linear equations in terms of d_k and a Lagrange multiplier ζ_k when the constraints in the active set are independent. That is, ζ_k is taken to satisfy

$$(A1.48) \quad \mathbf{A}_k^T \zeta_k = \nabla f(x_k) + B_k x_k$$

Initializing the algorithm requires an initial feasible point that can be found by solving the linear programming (LP) problem:

$$(A1.49) \quad \begin{aligned} & \min_{\xi \in \mathbb{R}, x \in \mathbb{R}^n} \quad \xi \\ & \text{subject to} \quad A_i x = -\hat{c}_i, \quad i = 1, \dots, m_{\text{eq}} \\ & \quad \quad \quad A_i x - \xi \leq -\hat{c}_i, \quad i = m_{\text{eq}} + 1, \dots, m \end{aligned}$$

which can be done by solving the under or over determined system of equations for the equality constraints. Once the feasible point is found, the initial search direction $d_{z,0}$ is found by solving the set of linear equations

$$(A1.50) \quad B_k d_{z,0} = -G_k$$

where G_k is the gradient to the objective function. At each iteration the updated Hessian of the Lagrangian is in the form of a positive definite quasi-Newton approximation given by:

$$(A1.51) \quad B_{k+1} = B_k + \frac{q_k r_k^T}{q_k^T s_k} - \frac{B_k s_k s_k^T B_k^T}{s_k^T B_k s_k}$$

where

$$s_k = x_{k+1} - x_k$$

$$q_k = \left[\nabla f(x_{k+1}) + \sum_{i=1}^m \zeta_i \cdot \nabla g_i(x_{k+1}) \right] - \left[\nabla f(x_k) + \sum_{i=1}^m \zeta_i \cdot \nabla g_i(x_k) \right]$$

Chapter III

A1.4 MHD Gravitational Solver

The integration scheme for the gravitational solver is demonstrated in Matsumoto and Hanawa [71] through solving Poisson's partial differential equation on a nested grid. The method is summarized below and begins with defining the Poisson equation as:

$$(A1.52) \quad \nabla^2 \Phi = 4\pi G \rho$$

where Φ is the gravitational potential. This can further be condensed as

$$(A1.53) \quad \nabla \cdot \mathbf{g} = -4\pi G \rho$$

where the gravity \mathbf{g} is defined by

$$\mathbf{g} = -\nabla \Phi$$

The Poisson equation is used to demonstrate how the quantities in the MHD equations are solved on the nested grid [71]. Begin by discretizing equation A1.53 by central differencing as:

$$(A1.54) \quad 4\pi G \rho_{i,j,k}^{(l)} = \frac{g_{x,i+1/2,j,k}^{(l)} - g_{x,i-1/2,j,k}^{(l)}}{2^{-l+1}h} + \frac{g_{y,i,j+1/2,k}^{(l)} - g_{y,i,j-1/2,k}^{(l)}}{2^{-l+1}h} + \frac{g_{z,i,j,k+1/2}^{(l)} - g_{z,i,j,k-1/2}^{(l)}}{2^{-l+1}h}$$

where

$$g_{x,i+1/2,j,k}^{(l)} = \frac{\Phi_{i+1,j,k}^{(l)} - \Phi_{i,j,k}^{(l)}}{2^{-l+1}h}, \quad g_{y,i,j+1/2,k}^{(l)} = \frac{\Phi_{i,j+1,k}^{(l)} - \Phi_{i,j,k}^{(l)}}{2^{-l+1}h}, \quad g_{z,i,j,k+1/2}^{(l)} = \frac{\Phi_{i,j,k+1}^{(l)} - \Phi_{i,j,k}^{(l)}}{2^{-l+1}h}$$

Equation A1.54 is valid for all grid cells interior to the grid level boundary and in order to remain consistent with the discrete version of Gauss's theorem,

$$(A1.55) \quad \sum_{\text{surface}} \mathbf{g} \cdot d\mathbf{S} = 4\pi G \sum_{\text{volume}} \rho d\mathbf{V}$$

the condition is set that the cell gravity on a coarse grid is equal to the average gravity evaluated on smaller grid cells, that is:

$$(A1.56) \quad g_{x,i+1/2,j,k}^{(l)} = \frac{1}{4} \left(g_{x,2i+3/2,2j,2k}^{(l+1)} + g_{x,2i+3/2,2j,2k+1}^{(l+1)} + g_{x,2i+3/2,2j+1,2k}^{(l+1)} + g_{x,2i+3/2,2j+1,2k+1}^{(l+1)} \right)$$

At the grid level boundaries however, equation A1.54 must be modified. Notice that when calculating the gravity at some surface, the gravitational potential at the neighbouring cell center is required. In order to get around this, the gravitational potential Φ is interpolated across the cell surface. To demonstrate, assume that $i = N/2 - 1$ (x boundary at the grid level l). The gravity at the surface of this cell is given as:

$$(A1.57) \quad g_{x,i+1/2,j,k}^{(l)} = \frac{\Phi_{i+3/2,j,k}^{*(l)} - \Phi_{i,j,k}^{(l)}}{(3/2)2^{-l+1}h}$$

where $\Phi_{i+3/2,j,k}^{*(l)}$ denotes the gravitational potential evaluated at the coarse cell surface along the x boundary (y and z boundaries are trivially modified)

$$(x, y, z) = (x_{(i+1)/2}^{(l-1)}, y_j^{(l)}, z_k^{(l)})$$

and is given as:

$$(A1.58) \quad \Phi_{i+3/2,j,k}^{*(l)} = \frac{3}{4} \Phi_{(i+1)/2,(j-1)/2,(k-1)/2}^{(l-1)} + \frac{1}{4} \Phi_{(i+1)/2,(j+1)/2,(k+1)/2}^{(l-1)}$$

which is valid for odd numbered j and k since indices must remain integer quantities (for $j = 14$, cannot have the $(14 + 1)/2 = 7.5^{\text{th}}$ cell). This of course is not required for even j and k and equation A1.58 is changed accordingly to maintain the cell indices as integer quantities. Equation A1.58 evaluates the gravitational potential along the diagonal of the coarse cell surface and uses the potential at both ends of the diagonal for linear interpolation. The gravity at the center of each cell is finally evaluated and is done so by taking the average at opposing cell surfaces:

$$(A1.59) \quad g_{x,i,j,k}^{(l)} = \frac{1}{2} (g_{x,i+1/2,j,k}^{(l)} + g_{x,i-1/2,j,k}^{(l)})$$

The difference equations are finally solved iteratively in conjunction with a full multigrid (FMG) scheme [72] to accelerate convergence. In essence, the multigrid iteration is used to provide a better initial guess when iterating over finer grids. That is, since the computation time on a coarse level is less, interpolating these outputs to a finer level provide a strong initial guess and thus numerical methods are expected to converge with fewer iterations. Matsumoto and Hanawa [71] introduce three temporary grids; one nested grid of $(N, h) = (2^{n-1}, 2h_0), (2^{n-2}, 2^2h_0), \dots, (2^2, 2^{n-2}h_0)$ and two uniform grids of 2^3 and 1 cell respectively covering the entire simulation grid. The output of one grid is effectively served as the initial guess to another. That is, beginning with an exact solution on the single celled grid, this is used as an initial value to determine the solution on the 2^3 celled grid and finally this output is used to obtain the solution on the nested grid at the coarsest level $(N, h) = (2^2, 2^{n-2}h_0)$ using successive over relaxation. The results are then interpolated using bilinear interpolation to serve as an initial guess for finer level $(N, h) = (2^3, 2^{n-3}h_0)$. At this level, red-black Gauss-Seidel iteration is used to perform the calculation. These operations are successively performed on finer and finer grids through a combination of interpolations, iterative calculations and restrictions. Restrictions are used to project the solution on a fine grid back to that on the coarse grid, Matsumoto and Hanawa [71] use a full weighted averaging. Figure 69 schematically demonstrates order of operation and the fully detailed algorithm is discussed with rigorous analysis in [71].

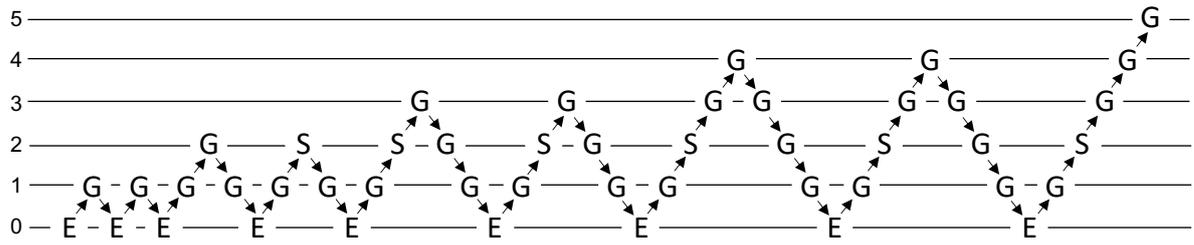


Figure 69: The simulation FMG scheme. The symbols **G**, **S** and **E** are given for Gauss-Seidel, Successive Over Relaxation and the exact solution respectively. The upward arrows denote interpolation and the downward ones denote weighted averaging. The numerical scale on the left hand side govern the grid level where 0 is representative of the coarsest level and 5 is representative of the finest level.

Chapter IV

A1.5 Eckmann Algorithm

To find the Lyapunov exponent of a given time series, the algorithm first introduced by Eckmann & Kamphorst [59] is implemented and can be described in 4 steps.

Step 1: Embed the time series

Given a time series $\mathbf{x}(t)$, select an embedding dimension d_E and embed $\mathbf{x}(t)$ in a d_E -dimensional orbit by defining:

$$\mathbf{x}_i(t) = [x_i(t), x_{i+1}(t), \dots, x_{i+d_E-1}(t)] \quad \text{for } i = 1, 2, \dots, N - d_E + 1$$

which effectively maps each point $x_i(t)$ of the time series to a vector $\mathbf{x}_i(t)$.

Step 2: Find the neighboring vectors

For each vector $\mathbf{x}_i(t)$ in the orbit, a suitable radius r_i must be found such that at least n_{\min} neighboring vectors lie within the Chebyshev distance r_i centered at $\mathbf{x}_i(t)$.

$$\begin{aligned} \|\mathbf{x}_i(t) - \mathbf{x}_j(t)\| &\leq r_i \\ \text{with} \\ \|\mathbf{x}_i(t) - \mathbf{x}_j(t)\| &= \max_{0 \leq \alpha \leq d_E-1} |x_{j+\alpha}(t) - x_{i+\alpha}(t)| \end{aligned}$$

where the vectors $\mathbf{x}_j(t)$ are the neighboring vectors.

Step 3: Compute the matrix \mathbf{T}

Define a matrix \mathbf{T}_i representative of the time evolution of how points are sent from the neighborhood of $\mathbf{x}_i(t)$ to the neighborhood of $\mathbf{x}_{i+1}(t)$. The matrix is of size $d_M \times d_M$ with matrix dimension $d_M \leq d_E$ corresponding to the time evolution from $\mathbf{x}_i(t)$ to $\mathbf{x}_{i+m}(t)$. Assume there is an integer $m \geq 1$ such that $d_E = (d_M - 1)m + 1$ where $\mathbf{x}_i(t)$ is associated with a corresponding d_M -dimensional vector

$$\mathbf{x}_i(t) = [x_i(t), x_{i+m}(t), x_{i+2m}(t), \dots, x_{i+(d_M-1)m}(t)].$$

For $m > 1$ the matrix \mathbf{T}_i can be found using the condition

$$\mathbf{T}_i(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \approx \mathbf{x}_{j+m}(t) - \mathbf{x}_{i+m}(t)$$

and takes the form

$$\mathbf{T}_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_{d_M} \end{bmatrix}.$$

Here, the coefficients a_k are found in the least squares sense by minimizing the quantity Δ

$$\Delta = \sum_{j \in S_i^E(r)} \left(\sum_{k=0}^{d_M-1} a_{k+1} [x_{j+km}(t) - x_{i+km}(t)] - [x_{j+d_M m}(t) - x_{i+d_M m}(t)] \right)^2$$

where $S_i^E(r)$ is the set of indices j that represent the number of neighbors of $\mathbf{x}_i(t)$ within the radius r_i . The quantity r_i is selected as follows:

- Count the number of neighbors x_i that correspond to increasing values of r .
- Stop when the number of neighbors exceeds $\min(2d_M, d_M + 4)$.
- If for this choice of r , the matrix \mathbf{T}_i does not have a previous minimal fixed rank, increase r_i again.

Step 4: Compute Λ via QR-Decomposition

After computing the sequence of matrices $\mathbf{T}_i, \mathbf{T}_{i+m}, \mathbf{T}_{i+2m}, \dots$, the Lyapunov exponents can be found through successive QR-decomposition of $\mathbf{T}_i \mathbf{Q}_{i-1}$ beginning with $i = 1$ having $\mathbf{Q}_0 = \mathbf{I}$ for

$$\mathbf{T}_{1+jm} \mathbf{Q}_j = \mathbf{Q}_{j+1} \mathbf{R}_{j+1}$$

The respective Lyapunov exponent Λ_k is found using the diagonal elements of the matrix \mathbf{R} :

$$\Lambda_k = \frac{1}{\Delta t K m} \sum_{j=0}^{K-1} \ln(\mathbf{R}_{j,(kk)})$$

where $K \leq (N - d_M m - 1)/m$ and N are the number of time steps in the time series and Δt is the respective step size. Finally, the maximum Lyapunov exponent is simply the largest value Λ_k from the generated sequence.

A1.6 OWA-ESN Outputs for Models 26 to 32

Optimized free running signal for models 26 to 32 excluding model 29 as demonstrated in the main body of section 20.7

Figures 70 to 75 demonstrate the OWA-ESN output for all remaining models (models 26 to 32 excluding 29) demonstrated in figure 52 where the implementation is done so exactly as demonstrated for model 29. As mentioned earlier, there are differences in the training, validation and free run segments given that each respective simulation model exhibits different chaotic periods that occupy different lengths of time. Therefore, the amount of data available for training and validation varies from one model to the next. This can hinder and effect the network’s performance, however one can still see strong performing outputs for at least one of the four possible deployments for each simulation model. Ideally like any form of neural networks, expectation is that the outputs improve with larger training and validation data available. All the remaining models implement 50 data points for testing and all the performance metrics on the standardized final output are summarized in table A1.1 below.

Table A1.1: Summary of the respective performance metrics on all remaining models for the network’s final output $\dot{M}(t)\sigma^{-1}$.

Low input scaling			High input scaling ($\varrho = 1$)		
Model	Readout	MSE	Model	Readout	MSE
Model 26	Ridge	1.48635	Model 26	Ridge	0.00159
	Pseudo	0.01675		Pseudo	0.00368
Model 27	Ridge	0.00035	Model 27	Ridge	0.00044
	Pseudo	1.41907e−5		Pseudo	2.74969e−6
Model 28	Ridge	8.45136e−5	Model 28	Ridge	0.00281
	Pseudo	7.97302e−7		Pseudo	5.52633e−7
Model 30	Ridge	0.06395	Model 30	Ridge	0.00020
	Pseudo	0.00312		Pseudo	0.00853
Model 31	Ridge	0.03446	Model 31	Ridge	0.01325
	Pseudo	0.03897		Pseudo	0.03945
Model 32	Ridge	0.00101	Model 32	Ridge	0.00014
	Pseudo	0.00707		Pseudo	0.00333

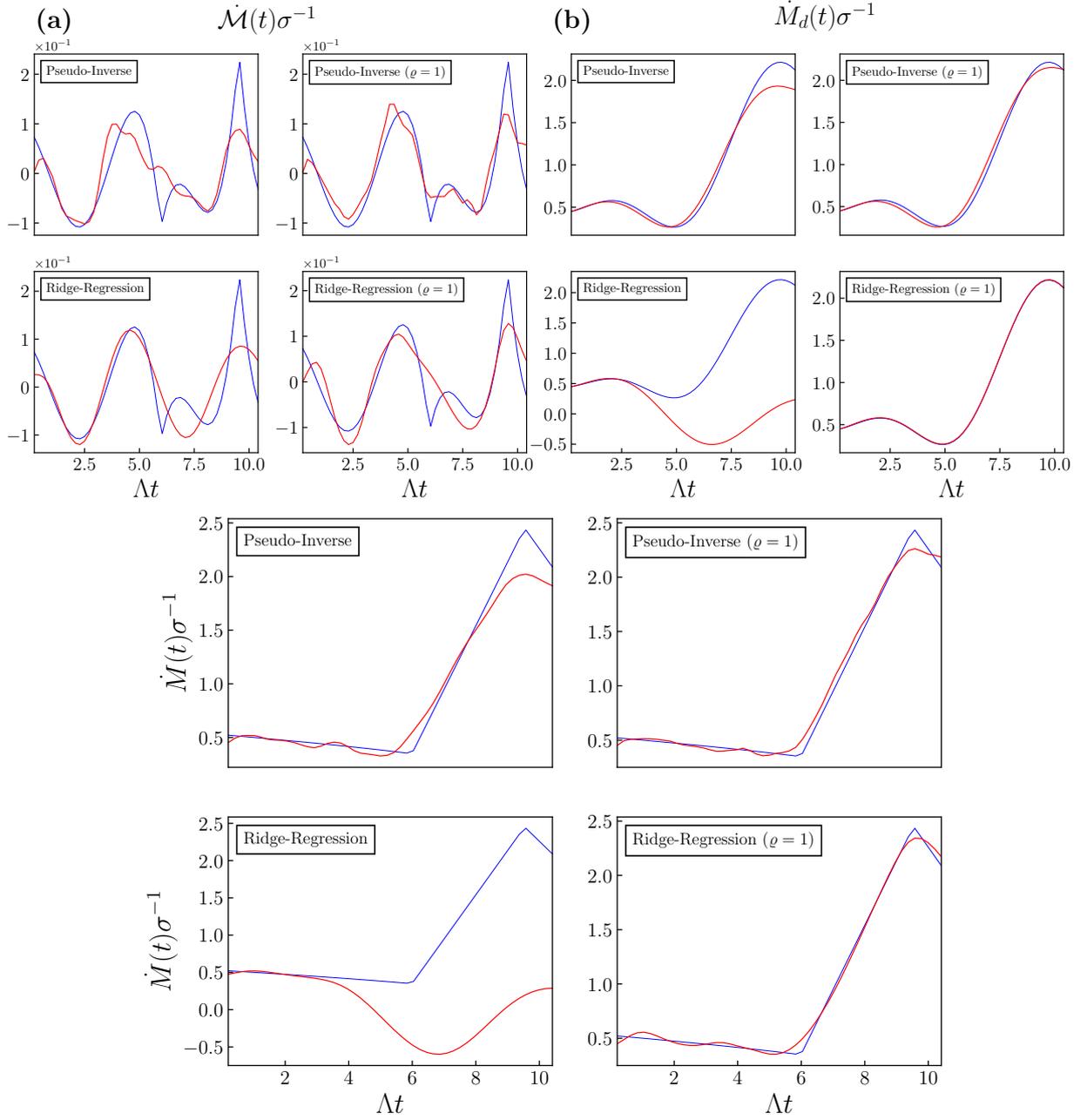


Figure 70: The optimized output signal (red) for model 26 overlaid with the free run segments (blue) of the standardized **(a)** fluctuating and **(b)** deterministic components given in terms of $\Lambda \equiv \Lambda_{\max} = 0.20821945$. The bottom subplot panel gives the final output $\dot{M}(t)$. Model 26 utilized a series of $M = 100$ reservoirs, had a standard deviation of $\sigma = 2.0231e-5$ and utilized 6450 data points where 3200 were used for training and 3200 for validating.

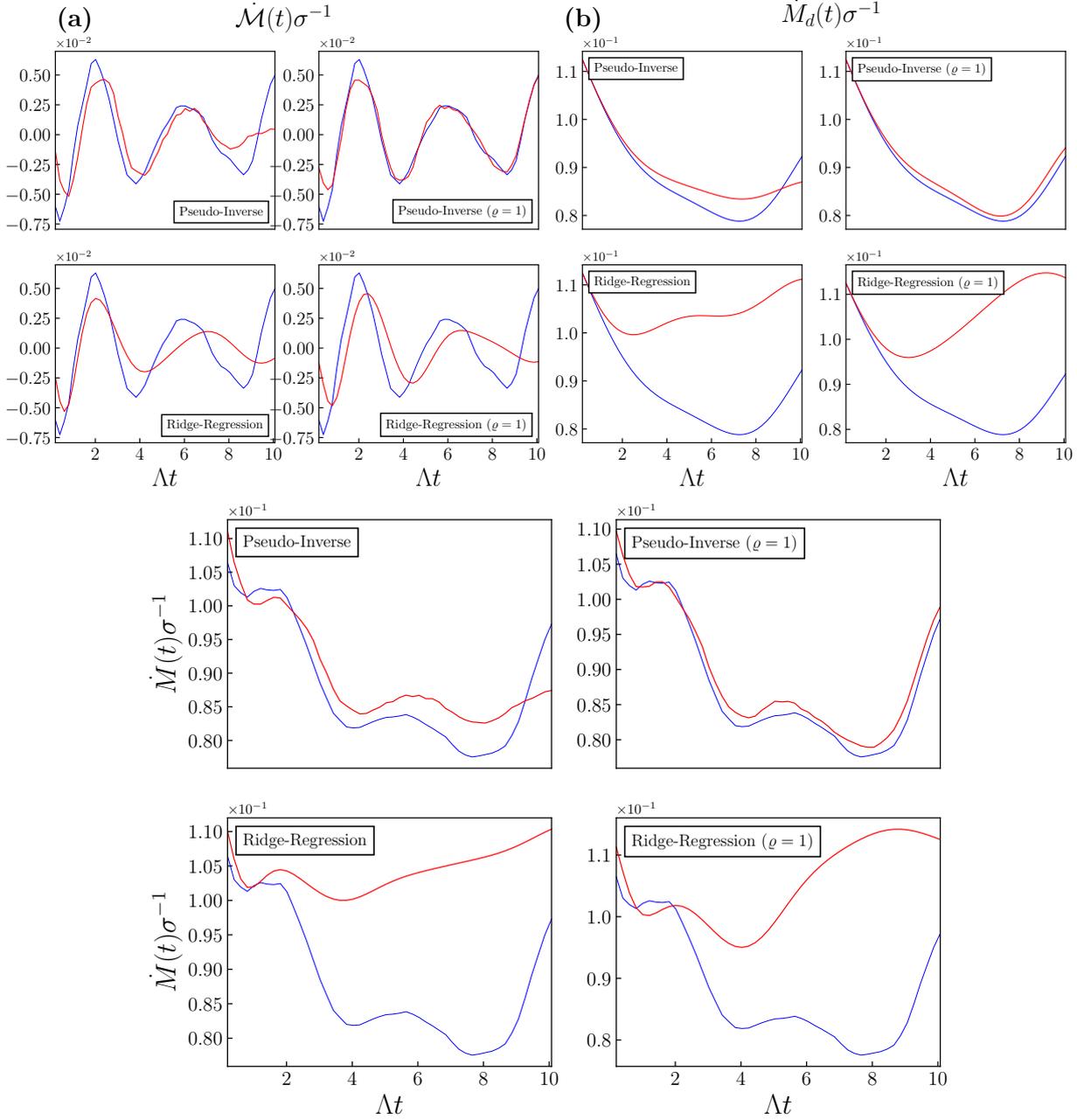


Figure 71: The optimized output signal (red) for model 27 overlaid with the free run segments (blue) of the standardized (a) fluctuating and (b) deterministic components given in terms of $\Lambda \equiv \Lambda_{\max} = 0.2014386$. The bottom subplot panel gives the final output $\dot{M}(t)$. Model 27 utilized a series of $M = 100$ reservoirs, had a standard deviation of $\sigma = 1.4145e-5$ and utilized 10050 data points where 5000 were used for training and 5000 for validating.

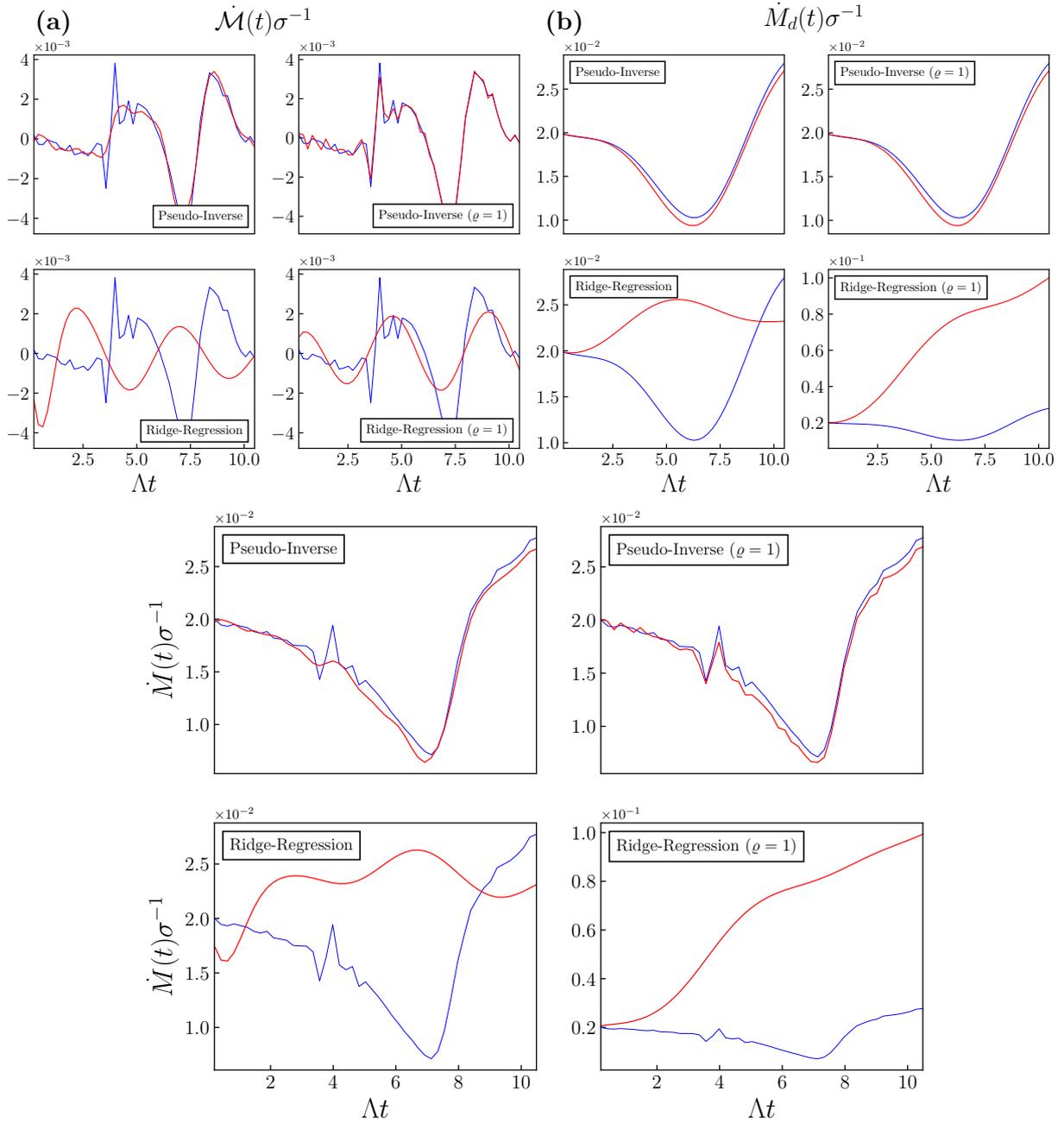


Figure 72: The optimized output signal (red) for model 28 overlaid with the free run segments (blue) of the standardized (a) fluctuating and (b) deterministic components given in terms of $\Lambda \equiv \Lambda_{\max} = 0.20973413$. The bottom subplot panel gives the final output $\dot{M}(t)$. Model 28 utilized a series of $M = 100$ reservoirs, had a standard deviation of $\sigma = 1.3449\text{e-}5$ and utilized 4050 data points where 2000 were used for training and 2000 for validating.

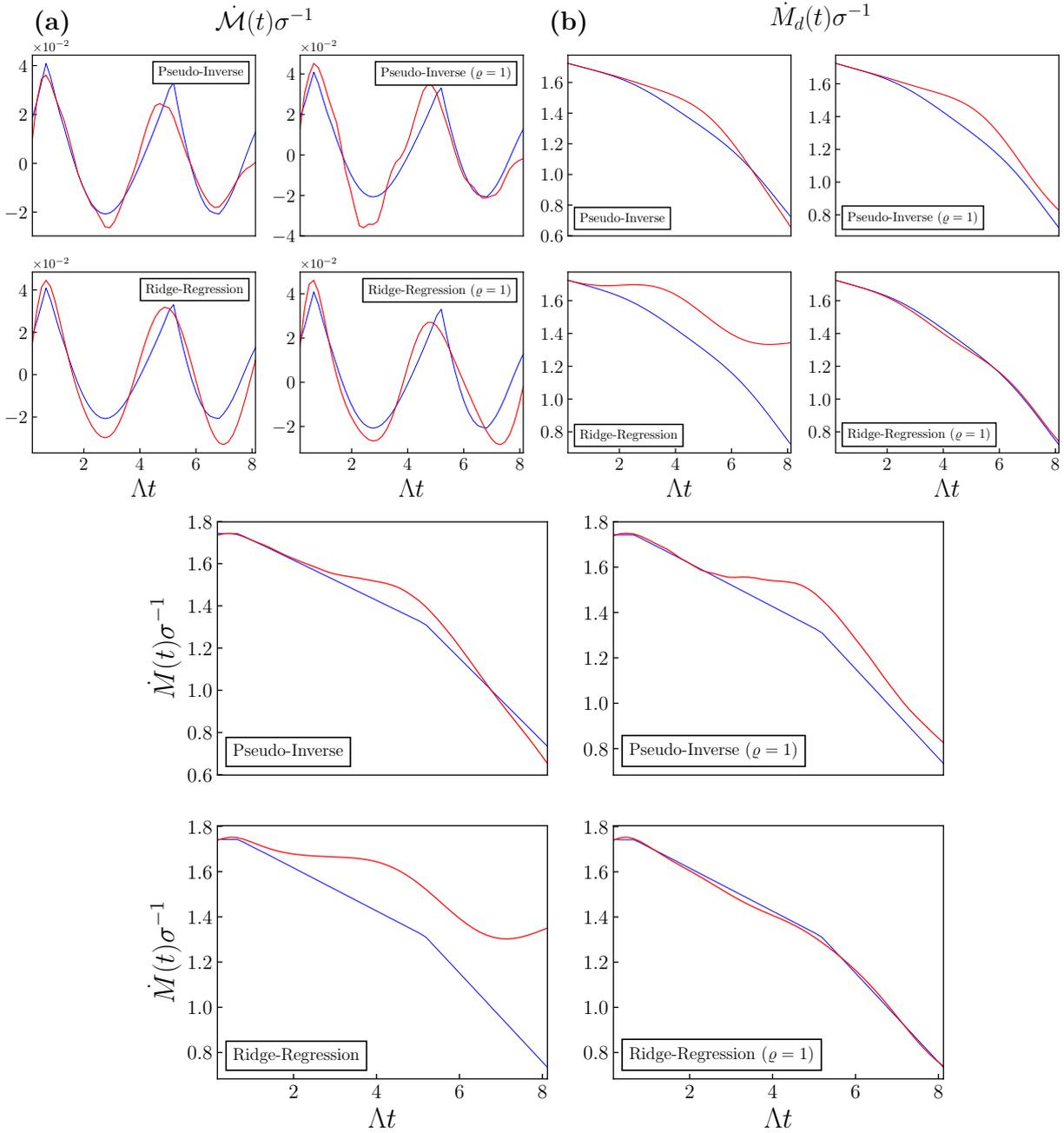


Figure 73: The optimized output signal (red) for model 30 overlaid with the free run segments (blue) of the standardized (a) fluctuating and (b) deterministic components given in terms of $\Lambda \equiv \Lambda_{\max} = 0.16244096$. The bottom subplot panel gives the final output $\dot{M}(t)$. Model 30 utilized a series of $M = 100$ reservoirs, had a standard deviation of $\sigma = 1.0350e-4$ and utilized 7850 data points where 3900 were used for training and 3900 for validating.

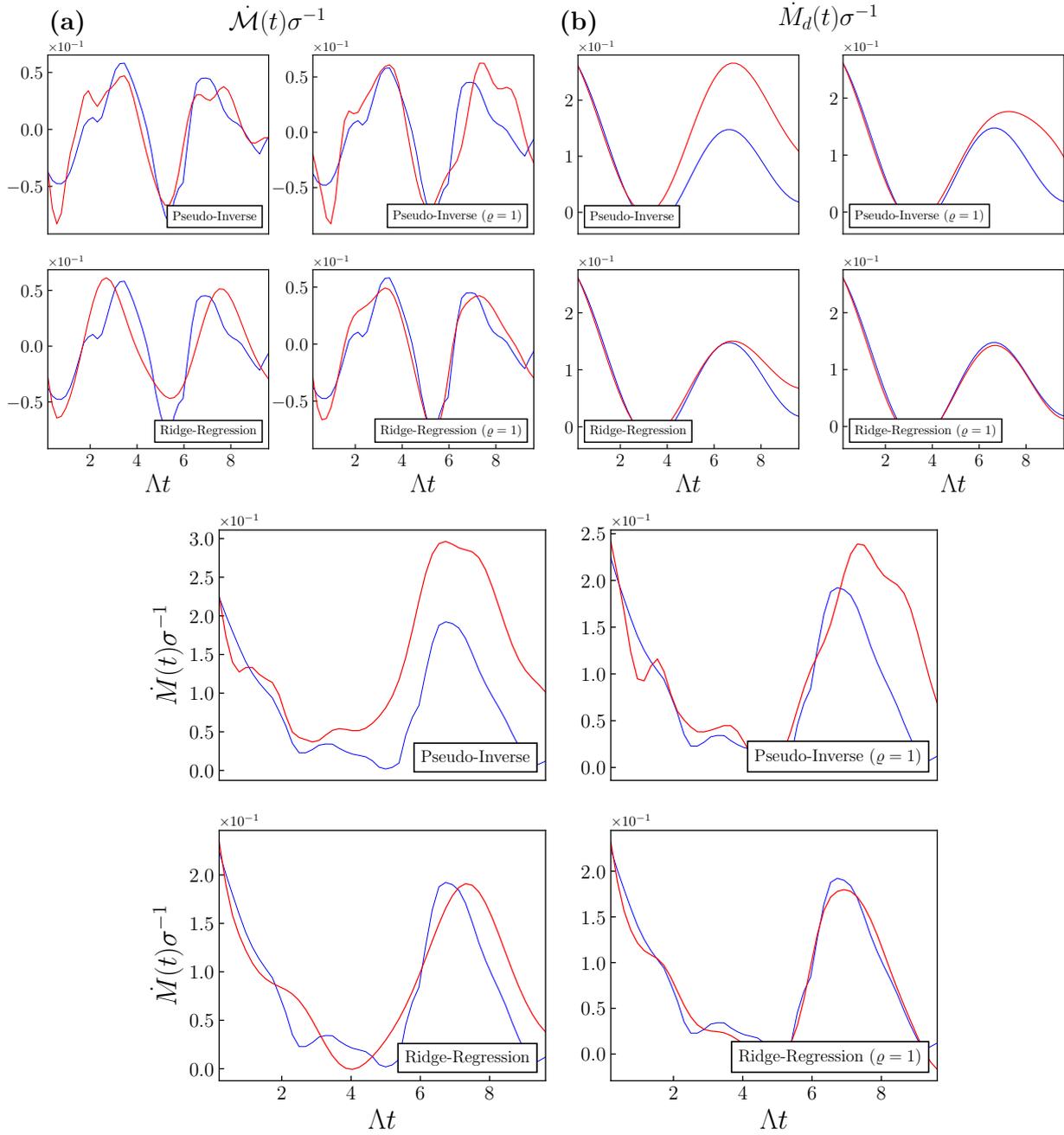


Figure 75: The optimized output signal (red) for model 32 overlaid with the free run segments (blue) of the standardized (a) fluctuating and (b) deterministic components given in terms of $\Lambda \equiv \Lambda_{\max} = 0.1922906$. The bottom subplot panel gives the final output $\dot{M}(t)$. Model 32 utilized a series of $M = 100$ reservoirs, had a standard deviation of $\sigma = 2.9740e-5$ and utilized 4950 data points where 2450 were used for training and 2450 for validating.