

UNIVERSITY OF CALIFORNIA

Los Angeles

Study and optimization of turbulence and transport in mirror configurations in the Large Plasma
Device

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Physics and Astronomy

by

Philip N. Travis

2024

ABSTRACT OF THE DISSERTATION

Study and optimization of turbulence and transport in mirror configurations in the Large Plasma
Device

by

Philip N. Travis

Doctor of Philosophy in Physics and Astronomy
University of California, Los Angeles, 2024

Professor Troy Carter, Chair

Abstract

The dissertation of Philip N. Travis is approved.

Paulo Alves

Christoph Niemann

Jacob Bortnik

Troy Carter, Committee Chair

University of California, Los Angeles

2024

TABLE OF CONTENTS

1	Introduction	1
1.1	Nuclear fusion: brief introduction	2
1.1.1	Mirror machines as a faster way to fusion power	2
1.2	The Large Plasma Device at UCLA	2
1.2.1	Diagnostics at the LAPD	2
1.2.2	Data acquisition	2
1.3	Analysis techniques	2
1.3.1	Spectral analysis	2
1.3.2	Correlation techniques	2
1.3.3	Machine learning	2
2	Turbulence and transport in mirror geometries in the Large Plasma Device	3
2.1	Introduction	4
2.2	The experiment and device configuration	6
2.2.1	The Large Plasma Device (LAPD)	6
2.2.2	Diagnostics	8
2.3	Mirror-induced changes	8
2.3.1	Profile modification	8
2.3.2	Reduced particle flux	12
2.3.3	Compensating for the Te profile	21
2.3.4	Drift waves	23
2.3.5	Turbulence modification	29

2.3.6	Magnetic fluctuations	30
2.4	2d Structure	36
2.5	Discussion	42
2.5.1	Lack of mirror-driven instabilities	42
2.5.2	Instabilities driving LAPD turbulence	46
2.5.3	Causes of particle flux reduction	47
2.5.4	Differences between DR1 and DR2	51
2.6	Conclusions and future work	54
3	Creating a randomized dataset for training machine learning models	56
3.1	Goal	56
3.2	The LAPD	56
3.3	Data bias	58
4	Optimizing mirror configurations in the LAPD using machine learning	63
4.1	Introduction	64
4.2	Data collection and processing	66
4.2.1	The Large Plasma Device (LAPD)	66
4.2.2	Data collection and processing of I_{sat} signals	66
4.2.3	Dataset distribution and bias	68
4.3	Model development, training	70
4.3.1	Training details	70
4.3.2	Validating the training pipeline	71
4.3.3	Baselines for mean-squared error	73

4.3.4	Effects of training set and model sizes	75
4.3.5	Improving performance with machine state flags	76
4.3.6	Learning rate scheduling	77
4.4	Uncertainty quantification	77
4.4.1	β -NLL loss	77
4.4.2	Cross-validation MSE	78
4.4.3	Model calibration via weight decay	81
4.5	Evaluating model performance	85
4.5.1	Checking geometrical intuition	85
4.5.2	Directly comparing prediction to measurement	87
4.5.3	Comparison with Thomson scattering	89
4.6	Effect of I_{sat} calibration	90
4.7	Inferring trends	92
4.8	Optimizing profiles	95
4.9	Discussion	101
4.9.1	Key achievements	101
4.9.2	Current limitations	102
4.9.3	Future directions	103
4.10	Conclusion	103
4.11	The open dataset and repository	104
5	Reconstructing missing diagnostics using energy-based models	105
5.1	Brief introduction to energy-based models (EBMs)	105
5.2	Building a 15 million shot dataset	105

5.3	Training the model	107
5.4	Unconditional sampling	107
5.5	Reconstructing missing diagnostics via conditional sampling	107
5.6	Inferring trends and outlook	107
6	Conclusions	108
6.1	Future Directions	108
A	0D mirror optimization	109
A.1	List of assumptions / conditions	110
A.2	User specified parameters	111
A.2.1	Simple mirror endplug	111
A.2.2	Tandem mirror	112
A.2.3	Engineering parameters	112
A.3	Fusion	113
A.3.1	Reactivity	113
A.3.2	Fusion power	114
A.4	General formulae	117
A.5	Radial particle transport	119
A.5.1	Classical diffusion	120
A.5.2	Bohm diffusion	120
A.5.3	Gyro-Bohm diffusion	121
A.5.4	ETG-driven transport	121
A.6	Mirror-specific derived quantities	122

A.6.1	Temperatures and confinement time in a beam-heated mirror from Egedal et al 2022 [EEF22]	122
A.6.2	Confinement time given by classical transport	124
A.6.3	End Cells/Plugs	125
A.6.4	Tandem mirror central cell	129
A.6.5	Overall power balance and plant power estimates	131
A.6.6	Instabilities	132
A.7	Costs and economics	133
A.7.1	Heating	133
A.7.2	Magnets	133
A.8	Optimization constraints	134
A.8.1	Midplane fields regularization via alpha particle confinement penalties . .	134
A.8.2	Kunal's suggestions	135
A.9	Optimizing mirror configurations	136
A.9.1	Gradient descent using SymPy and JAX	136
A.9.2	Example: optimizing Q in a simple mirror	137
B	Automated Langmuir sweep analysis using machine learning	141
C	Wisdom acquired	142
References	143

LIST OF FIGURES

2.1	Cartoon of the mirror turbulence experiment setup	7
2.2	Raw data processing steps	9
2.3	Shot-to-shot variation of I_{sat} profiles	11
2.4	Midplane I_{sat} profile	13
2.5	Line-integrated density	14
2.6	T_e from Langmuir sweeps at the midplane	15
2.7	Plasma potential and derived $\mathbf{E} \times \mathbf{B}$ velocity profiles	16
2.8	I_{sat} and B_{\perp} fluctuation power profiles	17
2.9	Cross-field, $\tilde{\mathbf{E}} \times \mathbf{B}$ particle flux	19
2.10	Phase and coherency of I_{sat} current and Vf near x_{PF}	20
2.11	Swept vs triple probe measurements	21
2.12	Triple probe Te and Te fluctuation profiles	22
2.13	Diffusivity relative to D_B	23
2.14	Diffusivity with Te compensation relative to D_B	24
2.15	I_{sat} gradients under varying profile smoothing methods	24
2.16	I_{sat} (density) fluctuation power	25
2.17	k_{\parallel} and coherency γ	26
2.18	Normalized density fluctuations vs ρ_s/L_n	27
2.19	I_{sat} scaled by L_n^2/n^2	28
2.20	Fluctuation power summed for each k_y	29
2.21	B_{\perp} fluctuation power spectra at various locations	31
2.22	B_{\perp} fluctuation power profiles for three regions	33

2.23	Summed fluctuation power of B_{\perp} in the core	34
2.24	B_{\perp} , flat field 500G vs 400G	35
2.25	B_{\perp} fluctuations at $x=0$ for different mirror lengths	36
2.26	B_z fluctuations in the core and x_{PF}	37
2.27	B_{\perp} fluctuation power profiles for low frequencies	38
2.28	B_z fluctuation power profiles for low frequencies	39
2.29	2d plane of the perpendicular magnetic field and the derived current density for $M = 1$	40
2.30	Perpendicular magnetic field and the derived current density for the $M = 1.9$ and $M = 2.68$ cases	41
2.31	Azimuthal mode number m amplitudes of I_{sat} fluctuations	43
2.32	k_y averaged about x_{PF}	44
2.33	Gradient scale length L_n	48
2.34	Particle flux: breakdown into components	49
2.35	I_{sat} decorrelation time	50
2.36	ExB shearing rate	51
2.37	Discharge power vs mirror ratio	52
2.38	I_{sat} profiles ($M=1$), DR1 vs DR2	53
2.39	Line integrated density, DR1 vs DR2	53
2.40	I_{sat} calibration factor over runs	54
3.1	Time-averaged I_{sat} distribution over shots	59
3.2	Distribution of probe x-coordinates in the dataset	60
4.1	size=12	67
4.2	size=12	68

4.3	Gas puff timings and example I_{sat} time series at three different z-axis locations from three different dataruns. Note that some discharges do not achieve steady state in I_{sat} .	69
4.4	size=12	70
4.5	size=12	72
4.6	size=12	74
4.7	size=12	79
4.8	Cross-validation test set error	80
4.9	Cross-validation training set error	81
4.10	Z-score distributions for the training and test sets for $\lambda = 0$	82
4.11	size=12	83
4.12	size=12	84
4.13	size=12	86
4.14	size=12	88
4.15	size=12	90
4.16	Comparison of I_{sat} predictions with Thomson scattering measurements	91
4.17	size=12	93
4.18	size=12	94
4.19	size=12	98
A.1	DT reactivities	114
A.2	DD reactivities	115
A.3	DT, DD, and D-He3 reactivity comparison	115
A.4	Simple mirror: cost function and gradient magnitude for each step	138
A.5	Simple mirror: magnetic fields for each step	139

A.6 Simple mirror: fusion power and Q for each step	139
A.7 Simple mirror: temperatures for each step	140

LIST OF TABLES

2.1	Magnetic mirror lengths and ratios	7
2.2	LAPD plasma parameters	10
2.3	x_c and x_{PF} locations for each mirror ratio	11
3.1	Data breakdown by class and dataset (percent)	61
4.1	Summary of test set losses for different training data and ensembles	75
4.2	Machine inputs and actuators for model inference	96
4.3	Machine inputs and actuators for optimized axial profiles	99

ACKNOWLEDGMENTS

I acknowledge peeps

VITA

- 2017 B.S. (Engineering Physics), University of Illinois at Urbana-Champaign
- 2018 Masters (Physics), University of California, Los Angeles
- 2024 Ph.D. (Plasma Physics), University of California, Los Angeles

PUBLICATIONS

Pubs

CHAPTER 1

Introduction

1.1 Nuclear fusion: brief introduction

1.1.1 Mirror machines as a faster way to fusion power

1.2 The Large Plasma Device at UCLA

1.2.1 Diagnostics at the LAPD

1.2.1.1 Langmuir probes: I_{sat} , sweeps, triple probes

1.2.1.2 The 288 GHz heterodyne interferometer

1.2.1.3 Thomson scattering

1.2.1.4 Fast framing camera

1.2.2 Data acquisition

1.3 Analysis techniques

1.3.1 Spectral analysis

1.3.2 Correlation techniques

1.3.3 Machine learning

1.3.3.1 Neural networks

CHAPTER 2

Turbulence and transport in mirror geometries in the Large Plasma Device

In this chapter we study turbulence and cross-field particle transport in LAPD mirror configurations. Mirror machines are once again rising in prominence as a candidate for commercial fusion reactors with the advent of highly-funded commercial ventures and high-field high-temperature superconducting magnets [EAB23, FAE24], so development of a functional understanding of cross-field transport in mirrors is imperative. Using the LAPD, multiple mirror ratios from $M = 1$ to $M = 2.68$ and three mirror-cell lengths from $L = 3.51\text{m}$ to $L = 10.86\text{m}$ were examined. Langmuir and magnetic probes were used to measure profiles of density, temperature, potential, and magnetic field. The fluctuation-driven $\tilde{E} \times B$ particle flux was calculated from these quantities. Two probe correlation techniques were used to infer wavenumbers and two-dimensional structure. Cross-field particle flux and density fluctuation power decreased with increased mirror ratio. Core density and temperatures remain similar with mirror ratio, but radial line-integrated density increased. The physical expansion of the plasma in the mirror cell by using a higher field in the source region may have led to reduced density fluctuation power through the increased gradient scale length. This increased scale length reduced the growth rate and saturation level of rotational interchange and drift-like instabilities. Despite the introduction of magnetic curvature, no evidence of mirror driven instabilities interchange, velocity space, or otherwise were observed. For curvature-induced interchange, many possible stabilization mechanisms were present, suppressing the visibility of the instability.

This chapter is largely a copy of my 2025 publication in the Journal of Plasma Physics titled

"Turbulence and transport in mirror geometries in the Large Plasma Device" [tra], with some additions and small changes

2.1 Introduction

Historically, mirror research has prioritized the main issues with mirror confinement: stabilizing the interchange instability, stabilizing velocity-space (loss-cone-driven) instabilities, and minimizing axial electron heat losses. Nevertheless cross-field transport remains an important topic in magnetic-confinement fusion reactor development, in both linear and toroidal geometries. Insight into edge-relevant turbulence, and its coupling to interchange and other mirror-driven instabilities, performed in a basic plasma science device may be useful for a mirror-based reactor. Although not at fusion-relevant core temperatures or densities, the Large Plasma Device (LAPD) operates at conditions similar to the edge of fusion devices and can provide insight into the physical processes in that region. A characterization of edge fluctuations has been undertaken, with emphasis on interpreting these fluctuations within the context of mirror.

Non-classical cross-field particle transport is often caused by low-frequency, large-amplitude fluctuations. These fluctuations are the result of various instabilities. One such process is the "universal" drift instability, which appears in the presence of a density gradient and finite resistivity. Drift wave turbulence and the effect on transport has been extensively studied in the past [Hor99, TFM09]. In the presence of sufficiently high rotation or sheared flow, rotational interchange and the Kelvin-Helmholtz instabilities also contribute or couple to these fluctuations.

Various gradient-, rotation-, and shear-driven instabilities (and suppression of such) have been studied previously in the LAPD experimentally [SCR12, SCR13, Sch13] and in simulations using BOUT, a 3d fluid turbulence code, and an eigenvalue solver [PUC10]. The LAPD has a sufficiently high spontaneous rotation rate that rotation-driven instabilities may be excited without artificial drive. Simulations using BOUT++ [FCU13] have also suggested that a rapidly growing nonlinear instability may dominate over all other linear instabilities.

Imposing a magnetic mirror configuration introduces magnetic curvature. The alignment of the curvature vector with a pressure gradient vector component causes the flute-like interchange instability if no stabilization mechanism is present. This interchange mode could couple to finite k_{\parallel} drift waves. The coupling of drift waves to curvature-induced interchange modes has been studied in toroidal devices such as TORPEX [PBD06, FLM06], where curvature was seen as the driving component for the unstable drift-interchange modes. Drift-like fluctuations have also been observed in the GAMMA-10 mirror [MII91, YMM10]. Flute-like modes and drift waves have been studied in other linear devices, such as Mirabelle [BGB05], where the appearance of flute-like modes or drift waves were controlled by varying the field and limiter diameter.

The rotational interchange and curvature interchange can both be flute-like modes. Rotational interchange (also called the “centrifugal instability”) is driven by the aligned centrifugal force and pressure gradient vectors, but curvature-driven interchange is instead driven by magnetic curvature and is typically referred to as simply the “flute” or “interchange” instability. Rotational interchange [Jas72] has been observed in the LAPD in the past [SCR12, Sch13], and the curvature-driven interchange instability has been observed in many other mirror machines [WV82, FWD83, Pos87].

Biassing or modifying the electrical connection of the plasma with the end wall has proven to be an important actuator in many mirror machines such as TMX-U [HBF84], GAMMA-10 [MII91], and GDT [BLZ03, BBS07, BBC10], and will be utilized on WHAM [EAB23]. Active biassing was not attempted in this study, but the intrinsic rotation and strong electrical connection to the source region may provide a useful analog for edge biassing in other mirror machines.

The LAPD exhibits a high degree of turbulence so it is difficult to identify the dispersion relation of the modes that are present. Nevertheless, the LAPD has good coverage of perpendicular spectra using correlation-plane techniques, and some measure of parallel spectra using the correlation between two axially-separated probes. A space-time spectral characterization of the many instabilities present in this low beta, moderate aspect ratio, gas-dynamic trap regime is attempted.

This goal of this study was to investigate the changes to turbulence and transport in LAPD mirror configurations. Of particular interest were the potential coupling of the interchange instability

with drift waves or other modes, and the effect of the mirror geometry on cross-field particle flux. Presented is a characterization of the observed modes and the effect of introducing a mirror geometry. This paper is organized as follows. Sec. 2.2 discusses the configuration of the LAPD and the diagnostics used. Sec. 2.3 covers the changes seen when imposing a magnetic mirror configuration on profiles, particle flux, drift waves, turbulence, and magnetic fluctuations. Sec. 2.4 explores the changes in 2d (x-y plane) structure. Sec. 4.9 discusses the active and expected instabilities and reasons for their modification. Sec. 4.10 summarizes the study and discusses the requirements for a deeper investigation.

2.2 The experiment and device configuration

2.2.1 The Large Plasma Device (LAPD)

In this study, the plasma was formed using an emissive, 72 cm diameter barium-oxide (BaO) cathode [GPL16a] (mapped to 60 cm in a flat field) and a 72 cm diameter, 50% transparent molybdenum anode that accelerate electrons across a configurable 40 – 70V potential; voltages of 60 and 63V were used in this study. The source has since been upgraded to a lanthanum hexaboride (LaB₆) cathode [QGP23a] that enables access to higher-density, higher-temperature regimes, but all the data in this study are from plasma formed by a BaO cathode.

The flexible magnetic geometry of the LAPD was used to construct a variety of magnetic mirror configurations. The discharge current, fill pressure, and other machine parameters were held constant. The typical plasma parameters observed in this study can be seen in table 2.2. Data in several mirror ratios and lengths were collected (see table 2.1) but emphasis is placed on the short cell because the highest mirror ratio possible ($M = 2.68$) with a 500 Gauss midplane field could be accessed and probes were able to be placed outside of the mirror cell. An overview of the axial magnetic field for the the short mirror configurations and probe locations can be seen in fig. 2.1. 2- or 3-cell mirror configurations were also explored but are not examined in this study. All results presented below are from the short mirror cell configuration unless otherwise specified.

Large Plasma Device

Side view

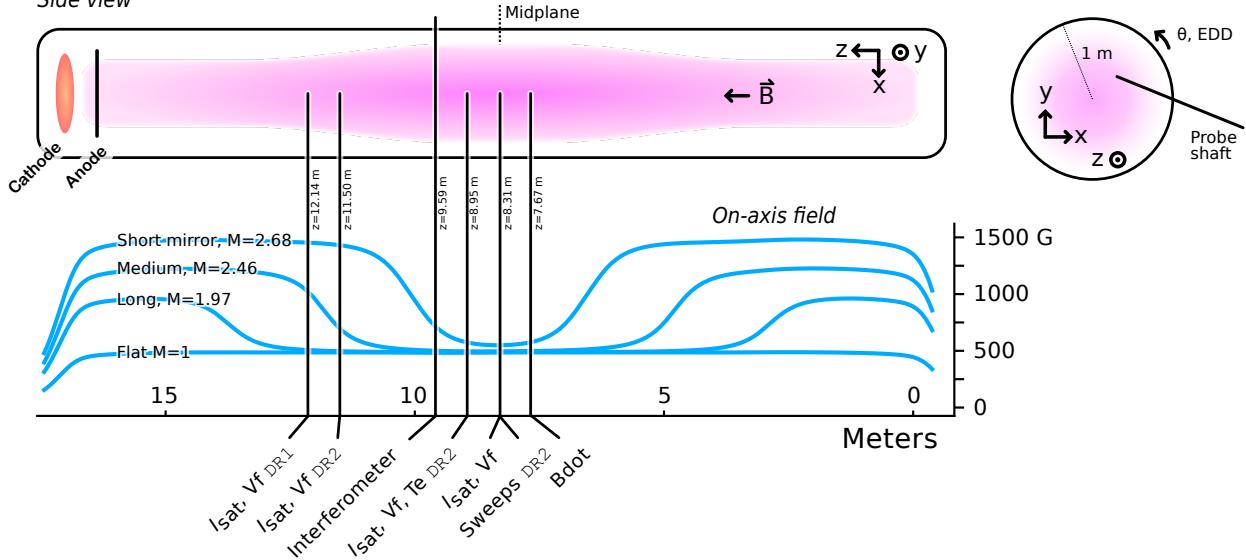


Figure 2.1: Cartoon of the Large Plasma Device and the coordinate system used. Only four of the eleven mirror configurations studied are plotted for clarity (mirrors of the same length have similar shapes and simply scale with mirror ratio). Diagnostic set varied by datarun; unlabeled diagnostics were used in both dataruns.

Mirror length	Mirror ratios (M)
Flat	1
3.51 m (short)	1.47 1.90 2.30 2.68
7.03 m (medium)	1.49 1.98 2.46
10.86 m (long)	1.47 1.97 2.44

Table 2.1: Magnetic mirror lengths and ratios. The lengths are measured where the curvature changes sign and the ratio is the maximum divided by the minimum. Approximately 3.5m must be added to the length if the good-curvature region is included. In the case of small asymmetries, the field strengths were averaged before calculation of the mirror ratio.

2.2.2 Diagnostics

All diagnostics were recorded with a effective sampling rate of 6.25 MHz (16-sample average at 100 MSPS) and a spatial resolution of 0.5 cm. When necessary, averaging over time is done in the approximate steady-state period of the plasma discharge (4.8 to 11.2 ms from the 1 kA trigger signal). Unless otherwise noted, all data presented will be from probes inside the mirror region ($z \approx 7\text{m}$). An example of a raw I_{sat} signal and processing steps can be seen in fig. 2.2. The raw signals are detrended by subtracting the mean across shots to obtain the fluctuations only. FFTs are then taken of these fluctuations for calculating power spectra and cross-correlated quantities. Frequencies above 200 kHz are dominated by electronics and instrumentation noise and thus are also ignored. Fluctuation power profiles can then be constructed.

The data presented were collected in two phases. The first phase ("datarun"), DR1, collected Langmuir probe (I_{sat} and V_f) and magnetic fluctuation ("Bdot") [EPC09] traces. 50 shots were taken at each position for every configuration. The second phase, DR2, was conducted with a similar set of diagnostics focused on temperature measurements (swept and triple probe) and 2d x-y structure. 15 shots were taken at each position, except for Langmuir sweeps with 64 shots. When appropriate, all data for each position were shot-averaged. " I_{sat} " will be used interchangeably with "density" and be presented with units of density (assuming a flat $T_e = 4.5\text{ eV}$ profile).

2.3 Mirror-induced changes

2.3.1 Profile modification

Because the field at the plasma source increases with M , the midplane plasma expands by a factor of \sqrt{M} . The physical locations of the peak fluctuation region – x_{PF} (maximum gradient) – and the cathode radius x_c can be seen in tab. 2.3. This expansion leads to broader plasma profiles and decreased core density but are similar in the core and at x_{PF} when magnetically-mapped to the cathode radius x_c as seen in fig. 2.4. Dips between the core ($x/x_c = 0$) and the peak fluctuation

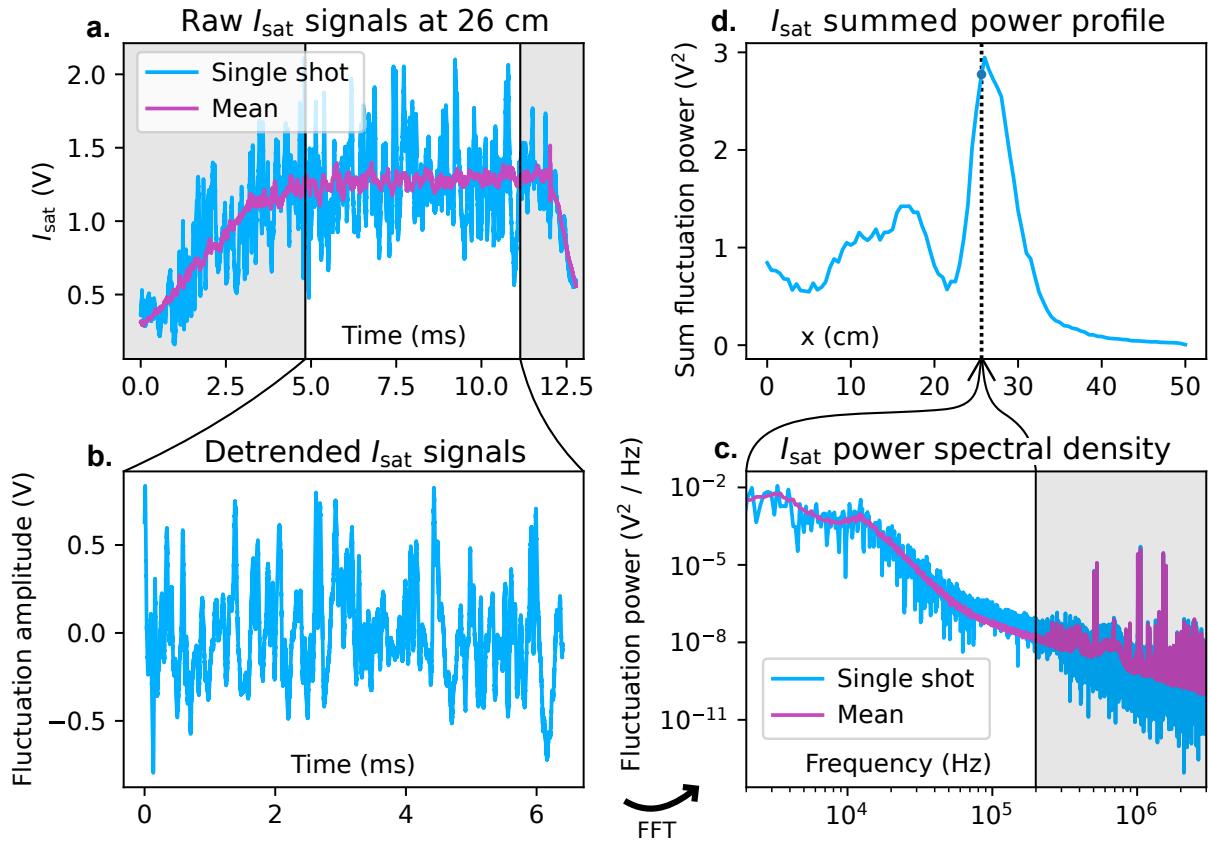


Figure 2.2: Raw data and basic processing steps for LAPD probe diagnostics as demonstrated by an I_{sat} trace from a DR1, $M = 1$ mirror at 26 cm. Data are truncated from 4.8 to 11.2 ms (a) and detrended (b). Power spectral density is calculated (c), and a power profiles can be constructed (d). The shaded regions are excluded from this analysis.

Cathode radius (M=1)	x_c	30	cm
Machine radius	R	50	cm
Plasma length	L	~ 17	m
Primary species		He-4 1+	
Electron-helium mass ratio		1.37×10^{-4}	
Neutral pressure		$6 - 20 \times 10^{-5}$	Torr
Quantity		Core	$x = x_{\text{PF}}$
			Unit
Density	n_e	1.25×10^{12}	0.6×10^{12} cm $^{-3}$
Ion temperature	T_i	~ 1	eV
Electron temperature	T_e	4	5
Beta (total)	β	9×10^{-4}	6×10^{-4}
Midplane magnetic field	B_{mid}	500	G
Plasma freq	Ω_{pe}	10	GHz
Ion cyclotron freq	Ω_{ci}	200	kHz
Electron cyclotron freq	Ω_{ce}	1.4	GHz
Debye length	λ_D	0.013	0.021
Electron skin depth	λ_e	30	43
Ion gyroradius	λ_{ci}	5.8	mm
Electron gyroradius	λ_{ce}	0.13	0.15
Ion thermal velocity	\bar{v}_i	6.94	km/s
Electron thermal velocity	\bar{v}_e	1190	1330
Sound speed	c_s	13.0	13.9
Alfvén speed	v_a	446 – 1140	–1620
Ion sound radius	ρ_s	65	69
Ion-ion collision freq	ν_{ii}	730	380
Electron-ion collision freq	ν_{ei}	6.77	2.59
Electron collision freq	ν_{ee}	9.57	3.66
Ion mean free path	$\lambda_{i,\text{mfp}}$	26	50
Electron mean free path	$\lambda_{e,\text{mfp}}$	175	512
Spitzer resistivity	η	192	$\mu\Omega \text{ m}$

Table 2.2: LAPD machine information and plasma parameters in the core and peak-fluctuation region ($x = x_{\text{PF}}$) at the midplane in this study. Dashed quantities are assumed to be identical to core quantities.

Mirror ratio	1	1.47	1.90	2.30	2.68
Scale factor	1	1.21	1.38	1.52	1.64
x_c (cm)	30	36	41	45	49
x_{PF} (cm)	26	32	36	40	43

Table 2.3: x_c and x_{PF} locations for each mirror ratio when scaled by the expected magnetic expansion.

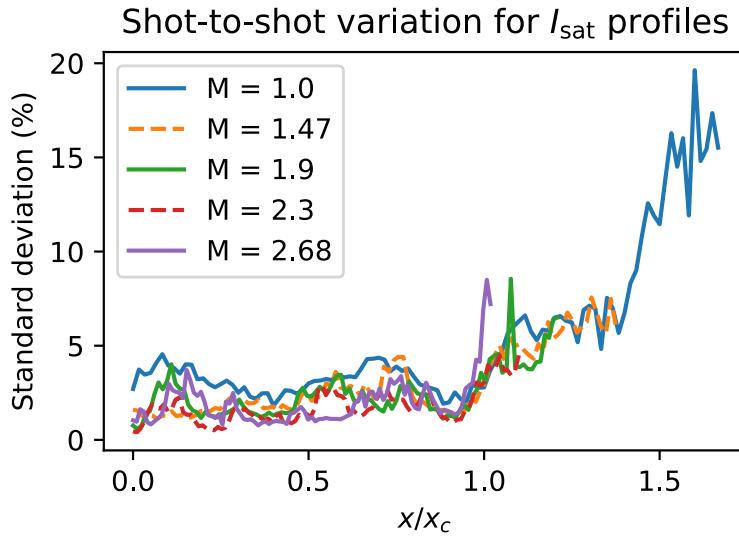


Figure 2.3: Shot-to-shot variation of I_{sat} profiles

region ($x = x_{PF}$) are seen, but fluctuation power from this region ($x/x_c = 0.5$ to 0.7) is not significant (fig. 2.8) so this region is not the focus of this study. The line-integrated density as measured by a 56 GHz heterodyne interferometer increases up to $\sim 35\%$ from the $M=1$ case of $\approx 8 \times 10^{13} \text{ cm}^{-2}$ (fig. 2.5) but does not increase past a mirror ratio of 2.3.

The error of on the I_{sat} profiles as represented by the standard deviation (scaled by the time-averaged profiles) can be seen in fig. 2.3. The error is relatively small and should not play a factor in our analysis – rarely are differences between quantities of the different mirror ratios that small.

Discharge power increases only slightly (3%) at higher mirror ratios suggesting negligible impact on density. Langmuir sweeps and triple probe measurements of T_e (DR2) show slightly (less than 25%) depressed core and slightly elevated edge T_e with increasing mirror ratio (fig. 2.6) but otherwise remains unaffected. The temperature affects I_{sat} measurements through the $\sqrt{T_e}$ term so small changes are insignificant. The low temperatures indicate that the plasma is collisional given the length scales of the system (as seen in table 2.2) and isotropic. Plasma potential decreases across the plasma (fig. 2.7) when the mirror ratio exceeds 1.9. This drop in plasma potential may be caused by the grounding of the anode to the wall, which should begin at $M = 1.93$ given the 72 cm anode and 100 cm vessel diameters. The reason for the local minimum in the M=2.68 is unknown. This potential profile creates a sheared $\mathbf{E} \times \mathbf{B}$ velocity profile (fig. 2.7) limited to 500 m/s in the core and exceeding ~ 3 km/s at the far edge. The flow does not exceed 4% of the sound speed (tab. 2.2) in the core or gradient ($x = x_{PF}$) region. The mirror ratio does not appear to significantly alter azimuthal flow. The floating potential (Vf) profile also exhibits similar behavior to the plasma potential (fig. 2.7), but is modified by the presence of primary electrons.

2.3.2 Reduced particle flux

The density fluctuation power peaks at the steepest gradient region ($x_{PF} = x/x_c \sim 0.88$) as expected as seen in fig. 2.8. x_{PF} occurs at nearly the same magnetically-mapped coordinate for each mirror ratio. These density fluctuations are a large driver of changes in the cross-field particle flux (eq. 2.1). Vf fluctuations also peak at the same location, but the total power across mirror ratios are similar and, relative to density fluctuations, much lower in the core. Core density fluctuations below 2 kHz are substantial in the core at lower mirror ratios, possibly caused by hollow profiles, nonuniform cathode emissivity, or probe perturbations, but are outside the scope of this study.

A spectral decomposition technique is used to calculate the time-averaged particle flux [Pow74] as seen in fig. 2.9:

$$\Gamma_{\tilde{E} \times B} = \langle \tilde{n} \tilde{v} \rangle = \frac{2}{B} \int_0^{\infty} k(\omega) \gamma_{n\phi}(\omega) \sin(\alpha_{n\phi}) \sqrt{P_{nn}(\omega) P_{\phi\phi}(\omega)} d\omega \quad (2.1)$$

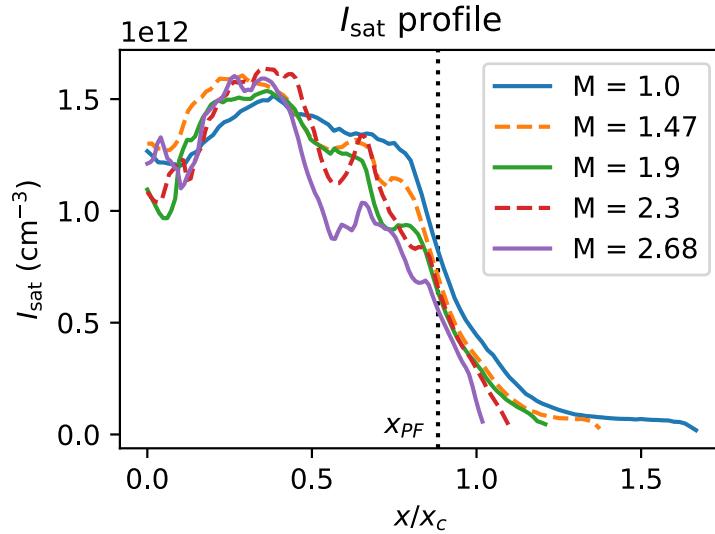


Figure 2.4: Midplane I_{sat} profile, shot-averaged and time-averaged from 4.8 to 11.2 ms (assumed of $T_e = 4.5$ eV based on triple probe and Langmuir sweep measurements). Effective area was calibrated using a nearby interferometer. Profile shape remains similar in the core and gradient region when mapped to the cathode radius x_c . The dips in profiles at higher M below $x = x_{PF}$ are of unknown origin and are not the focus of this study. Shot-to-shot variation is less than 5% for $x \leq 0.95x_c$ and less than 9% for $x \leq 1.4x_c$ for all cases.

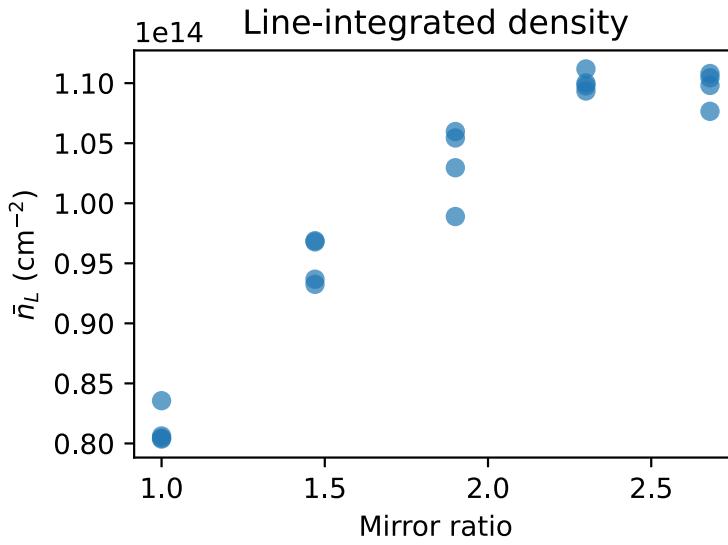


Figure 2.5: Line-integrated density as measured by a 56 GHz heterodyne interferometer as a function of mirror ratio, taken from four discharges for each mirror configuration. Density increases up to a mirror ratio of 2.3 where it appears to level off. The interferometer is located in the mirror cell bad-curvature region at 9.59m, 1.3m closer to the cathode from the midplane.

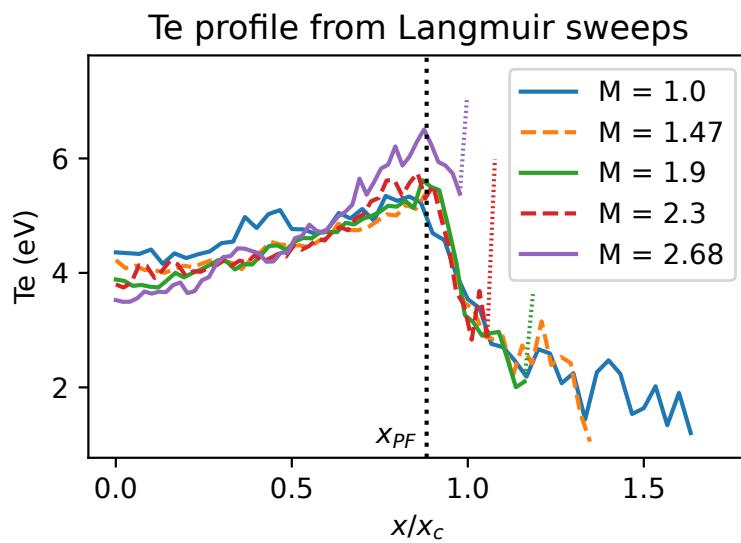


Figure 2.6: T_e from Langmuir sweeps (DR2) at the midplane. Triple probe results are nearly identical. The increased temperatures directly at the plasma edge, indicated by dotted portions of the curves, are likely artifacts caused by sheath expansion in lower densities. Changes in mirror ratio lead to at most 25% change in T_e . The plasma is collisional and isotropic.

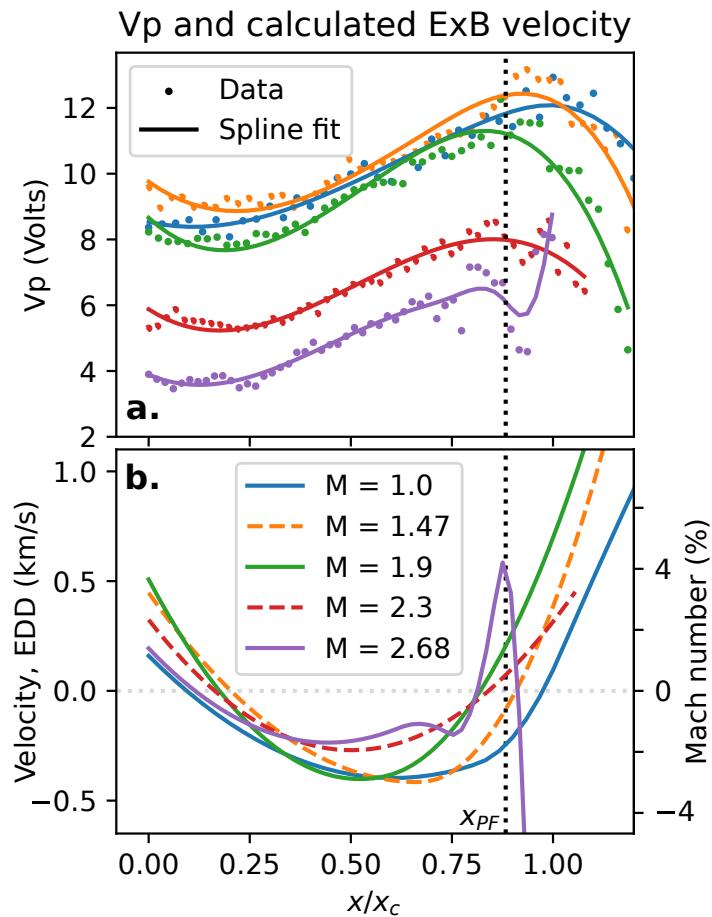


Figure 2.7: Plasma potential (a) and derived $\mathbf{E} \times \mathbf{B}$ velocity profiles (b) from Langmuir sweeps at the midplane. $x/x_c > 1.2$ has been excluded from the graph for greater clarity in the core and gradient region. The electric field was calculated by taking the gradient of the spline-smoothed plasma potential profile. The Mach number (in percent) is calculated using the approximate sound speed evaluated at $x = x_{PF}$ (tab. 2.2). The overall structure of the flows does not appreciably change when mirror ratio is varied.

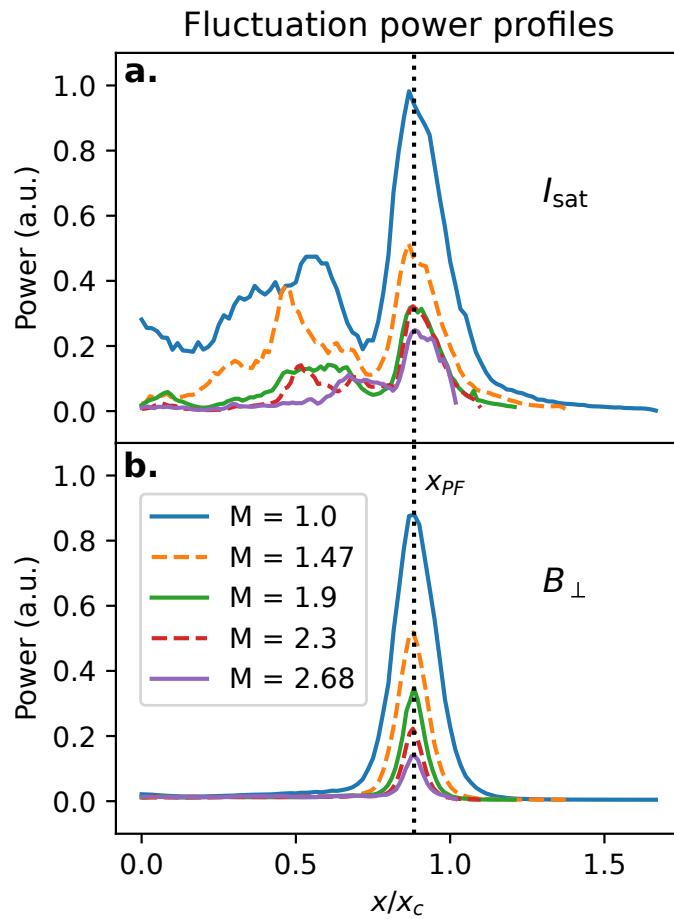


Figure 2.8: I_{sat} (a) and B_{\perp} (b) fluctuation power profiles for signals 2 kHz and up at $z=8.3\text{m}$ (midplane) and $z=7.7\text{m}$, respectively. The lower frequency components in I_{sat} are associated with bulk profile evolution, dominate the core region, and are not the focus of this study.

where k is the azimuthal wavenumber, γ is the coherency, α is the cross-phase, and P the power spectrum. This method is more robust than the naive time-integration of $n(t)\tilde{E}(t)$ because it accounts for the coherency of the density-potential fluctuations. This representation also enables inspection of each contributing term in the event of surprising or problematic results. A plot of the I_{sat} -Vf phase can be seen in fig. 2.10. The flattened particle flux in the core is likely caused by primary electrons emitted by the cathode. These electrons have long mean free paths (greater than a few meters) and sample fluctuations along the length of the machine, mixing the phases of these fluctuations. Since floating potential is set by the hotter electron population, the measured Vf fluctuations are no longer related to the local plasma potential fluctuations of a wave by bulk T_e [CM09]. These primary electrons have a significant effect in the core within the region mapped to the cathode $x \lesssim x_c$. I_{sat} fluctuations are not affected.

Azimuthal wave number is measured by two Vf probe tips 0.5 cm apart. This wavenumber estimation technique yields good agreement with correlation plane measurements (fig. 2.31). Note that \tilde{E} is not directly measured – it is instead calculated through the $k(\omega)\sqrt{P_{\phi\phi}(\omega)}$ terms. The $\tilde{E} \times B$ particle flux clearly decreases with mirror ratio; most of this decrease is attributed to the decrease in density fluctuation power. The particle flux for each mirror ratio was normalized to the $M = 1$ case via the plasma circumference to compensate for the increased plasma surface area at the same magnetically-mapped coordinate x/x_c . This particle flux is on the order of Bohm diffusion $D_B = \frac{1}{16} \frac{T_e}{B} \approx 6.25 \text{ m}^2 \text{s}^{-1}$ as observed in other transport studies [MCT07].

T_e profiles and fluctuations may affect particle fluxes but measurements of both were not taken in the same datarun; nevertheless, a quantification of the effect of T_e on particle flux is attempted. T_e fluctuations affect I_{sat} -based density measurements through the $T_e^{-1/2}$ term, and triple probe and Langmuir sweep T_e measurements suggest that temperature gradients have a negligible impact. A naive incorporation of temperature fluctuation data from DR2 into particle fluxes from DR1 suggest that cross-field particle flux may be underestimated by up to 50% via the I_{sat} temperature term, but the trend and relative fluxes across mirror ratios remain unchanged. Such a naive incorporation should be treated with suspicion because of the sensitive nature of the flux with respect to the

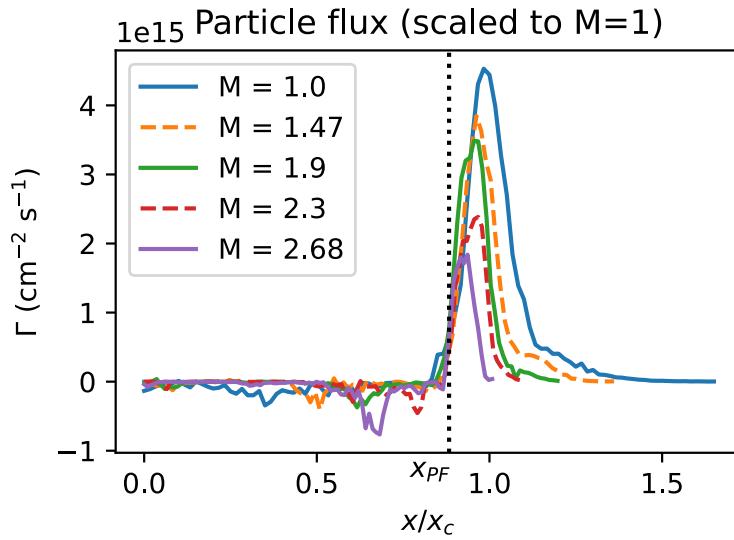


Figure 2.9: Cross-field, $\tilde{E} \times B$ fluctuation-based particle flux (calculated using eq. 2.1) with respect to mirror ratio. A monotonic decrease in particle flux is observed with increasing mirror ratio at the midplane. Particle flux is normalized by plasma circumference to the $M = 1$ case to account for the geometry-induced decrease in particle flux caused by a larger-diameter plasma.

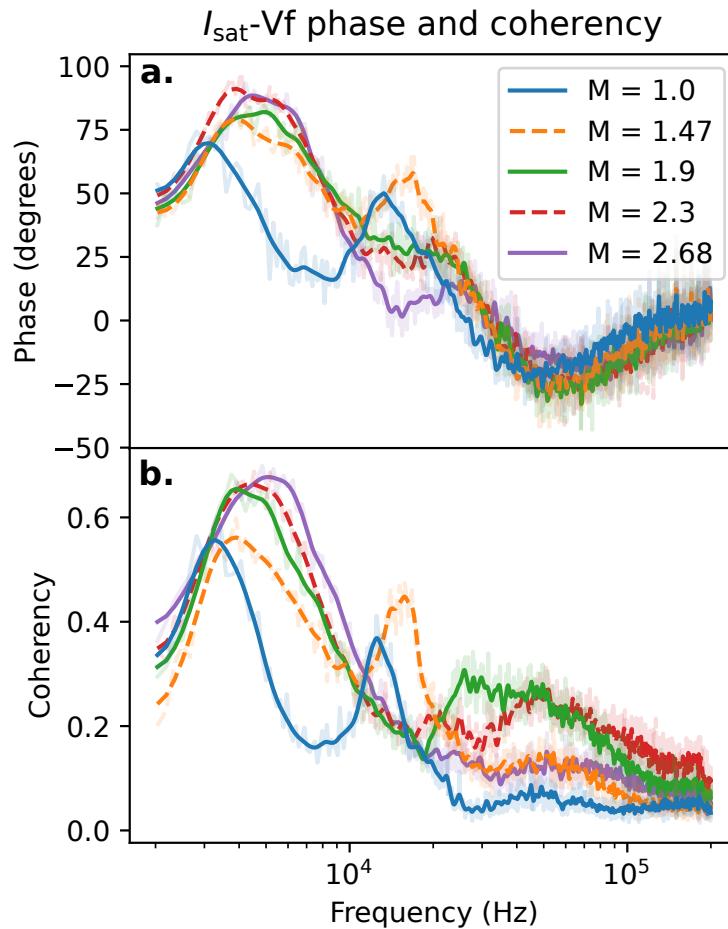


Figure 2.10: Phase (a) and coherency (b) of I_{sat} current and Vf near x_{PF} at the midplane, smoothed. Positive phase means I_{sat} leads Vf. Peaks in coherency occur between 3-5 kHz and at the drift-Alvén wave peaks between 12 and 25 kHz. These coherency peaks tend to have larger phase shifts than other nearby frequencies.

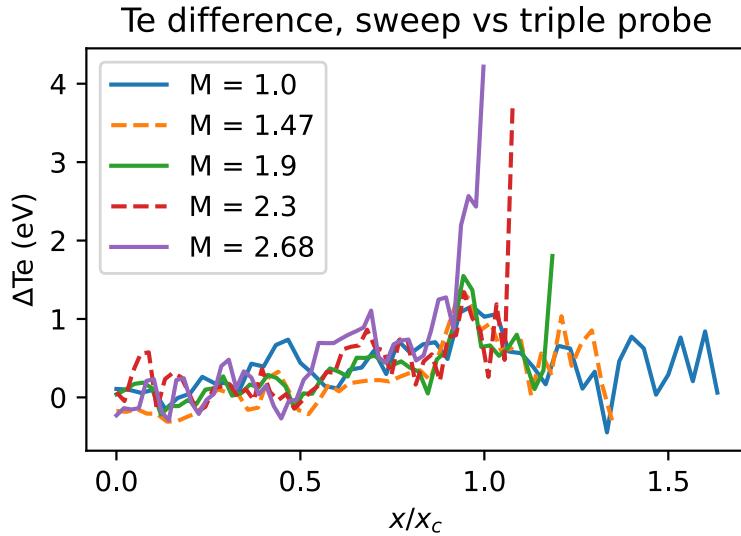


Figure 2.11: Difference between swept and triple probe temperature measurements. The triple probe appears to slightly underestimate the temperature and temperature gradient.

gradient and the differences in profiles between DR1 and DR2. These difference in profiles made be caused by cathode condition, deposits on the anode, or a different gas mix and are difficult to account for.

2.3.3 Compensating for the Te profile

Electron temperature (Te) compensation for the I_{sat} measurement can be done in several ways. One way is to account for the average temperature (i.e., steady state) when calculating the density from I_{sat} . Te can be gathered from triple probe or swept measurements. Triple probe measurements are generally less reliable than swept probe measurements. The difference between swept and triple probe Te measurements can be seen in fig. 2.11. The two techniques have roughly good agreement, though the triple probe appears to slightly underestimate the temperature. The spikes in the edge are likely from sheath expansion of the probe in the swept measurements (see fig. 2.6).

Te fluctuations can affect I_{sat} fluctuation measurements through the $\sqrt{\text{Te}}$ term. In this case, Te

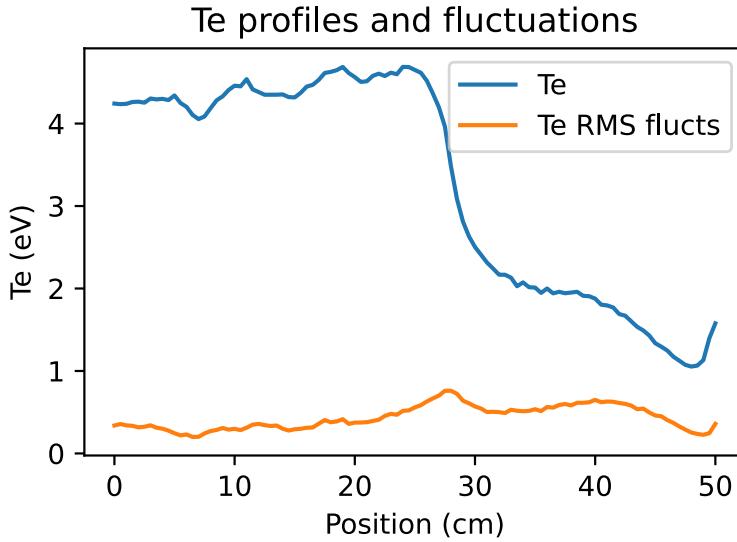


Figure 2.12: Te and Te fluctuation profiles from the triple probe. RMS electron temperature fluctuations are not particularly large.

measurements are difficult to compensate for in DR1 because of the changes in profiles between DR1 and DR2, so the Te fluctuations were included by finding the ratio in DR2 of I_{sat} fluctuations before and after including these Te fluctuations. This ratio was then applied to DR1. The issue of mismatched profiles still persists but this method allows for changes in fluctuation power between the two datarun sets. In general, $\tilde{\text{Te}}/\text{Te}$ fluctuations are at most than 30% (near the edge), and much lower in the core seen in fig. 2.12.

This Te compensation becomes particularly important when calculating the I_{sat} profile gradients which is needed when calculating the diffusivity. A calculation of the diffusivity scaled to the Bohm diffusivity $D_B = \frac{1}{16} \frac{T}{eB}$ can be seen in fig. 2.13. This calculation uses the particle flux calculated earlier (in the paper) and tanh fit on the density profile for a density smooth gradient. In general, mirror ratios higher than two have a lower diffusivity. When the particle flux is compensated for Te fluctuations, the temperature profile used in for the Bohm diffusion coefficient, and the density profile is smoothed convoluting a $\sigma = 2$ cm gaussian, the diffusion coefficient relative to D_B are roughly 2.5 times greater, seen in fig. 2.14. The trend, however, remains relatively

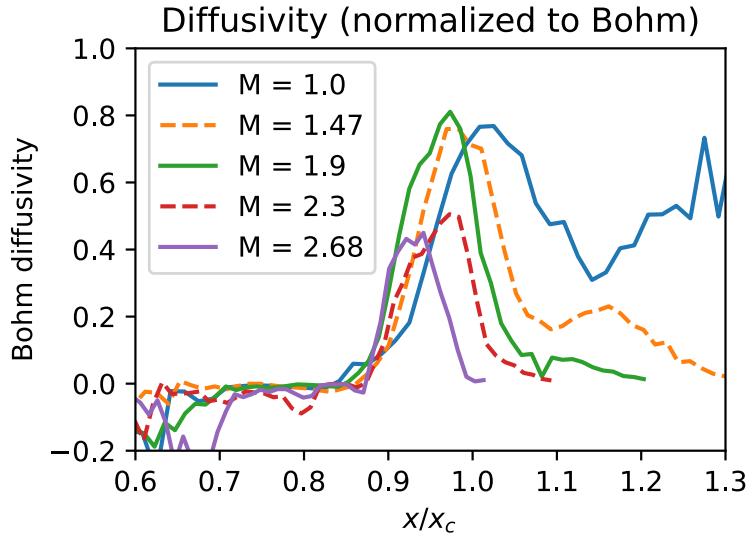


Figure 2.13: Diffusivity relative to D_B using a tanh fit for the density profile and the particle flux measurement assuming a constant T_e of 4.5 eV across the profile.

the same: higher mirror ratios tend to have a lower diffusivity. The impact of different profile smoothing methods on the density gradient can be seen in fig. 2.15.

2.3.4 Drift waves

The I_{sat} fluctuation power spectra in the region of peak power $x \sim x_{\text{PF}}$, also where the density gradient is strongest, can be seen in fig. 2.16. Notably, the fluctuation peaks shift to higher frequencies and decrease in total fluctuation power. The shift in frequency may be the Doppler shift caused by the change $\mathbf{E} \times \mathbf{B}$ plasma rotation seen in fig. 2.7 at the location $x/x_c \approx x_{\text{PF}}$. The shift in frequency is somewhat smaller than what would be expected from the field line-averaged increase in Alfvén speed at the longest possible wavelength. The phase angle of I_{sat} and V_f provides insight into the nature of the driving instability. Including a nonzero resistivity η in the drift wave leads to a small phase shift δ between density and potential. This phase shift δ in a collisional plasmas is on the order of $\delta \approx \omega v_e / k_{\parallel}^2 \bar{v}_e^2$ [Hor99]. Estimating this quantity using measured and

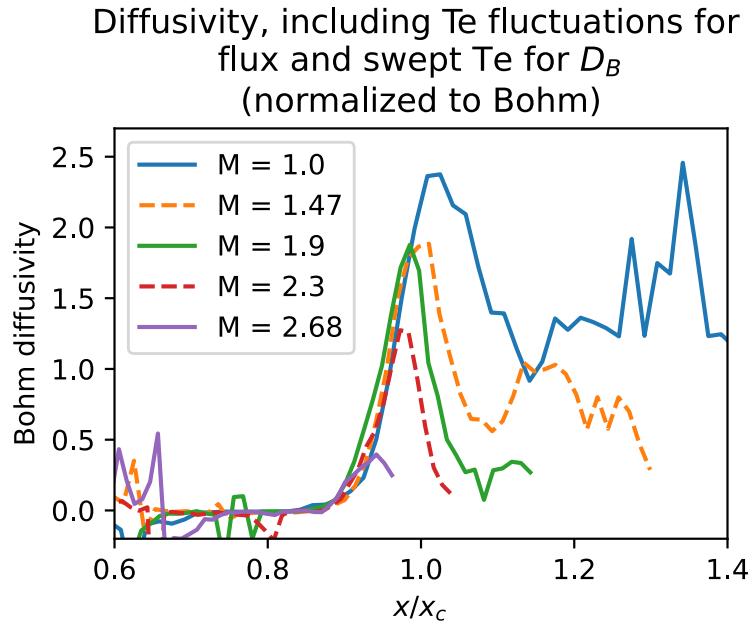


Figure 2.14: Diffusivity with Te compensation relative to D_B . The particle flux is compensated for Te fluctuations, and the swept-probe temperature profile is used for Te. The diffusivity is around 2.5 times higher than without compensation, but the trend remains similar.

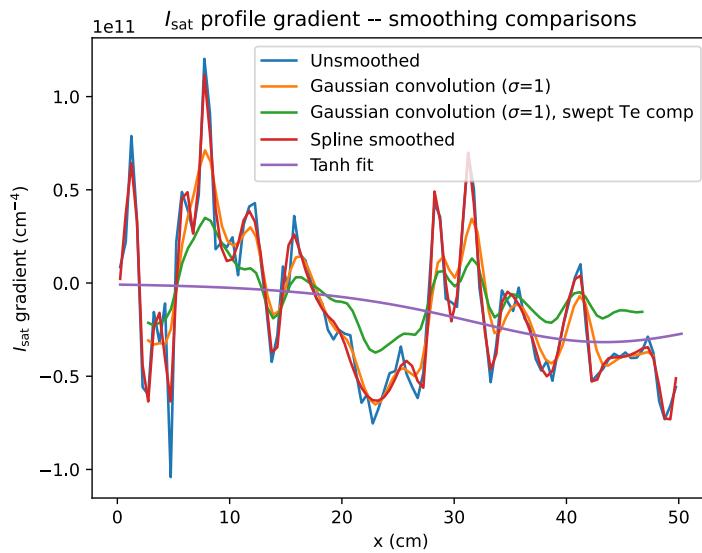


Figure 2.15: I_{sat} gradients under varying profile smoothing methods

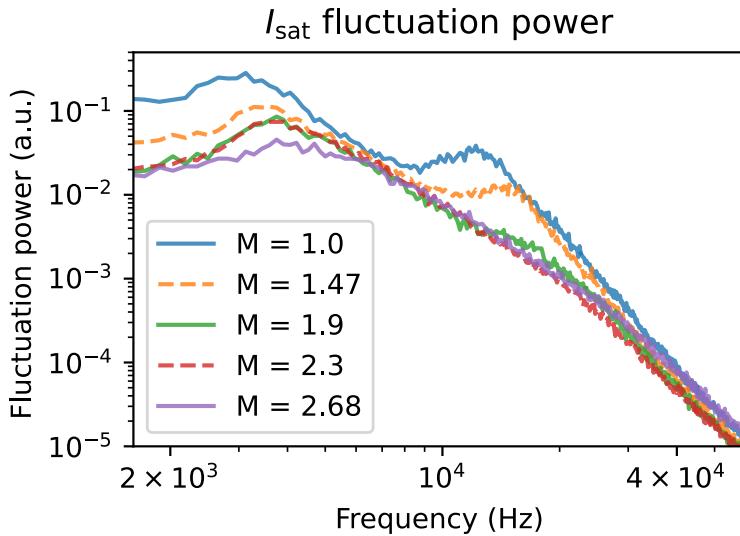


Figure 2.16: I_{sat} (density) fluctuation power averaged over a 1 cm region around x_{PF} at the mid-plane. The fluctuation power is largely featureless below 2 kHz and beyond 40 kHz aside from electronics noise.

typical values ($k_{\parallel} = 0.18 \text{ rad/m}$, $\bar{v}_e = 1300 \text{ km/s}$, $v_e = 3.7 \text{ MHz}$, $\omega = 12 \text{ kHz}$) yields a substantial phase shift of $\delta \approx 46^\circ$, which roughly agrees with the phase shifts in fig. 2.10, though the implied increased phase shift at higher frequencies does not agree with measurements. As seen in fig. 2.10, the phase shift between I_{sat} and Vf fluctuations are larger below 10 kHz, implying the presence of additional modes beyond or significant modification of resistive drift wave fluctuations. The phase difference between two Vf probes, 3.83 m apart, was used to calculate the parallel wavelength $\frac{2\pi}{\lambda} = k_{\parallel} = \phi_{\text{Vf1}, \text{Vf2}}/\Delta z$ assuming the wavelengths are greater than 7.66 m. The two probes mapped to the same field line only in the $M = 1$ configuration, so parallel wavenumbers are available only for the flat case. Parallel wavenumbers are theoretically calculable from 2d correlation planes but the coherency dropped dramatically when a mirror geometry was introduced. A 34 m wavelength mode likely contributes to the measured k_{\parallel} from 3 to $\gtrsim 10 \text{ kHz}$ (fig 2.17). Drift waves are long-wavelength modes so coherent density and potential fluctuations along the flux tube are expected. The coherency is a measure of similarity of the spectral content of two signals, in this case Vf

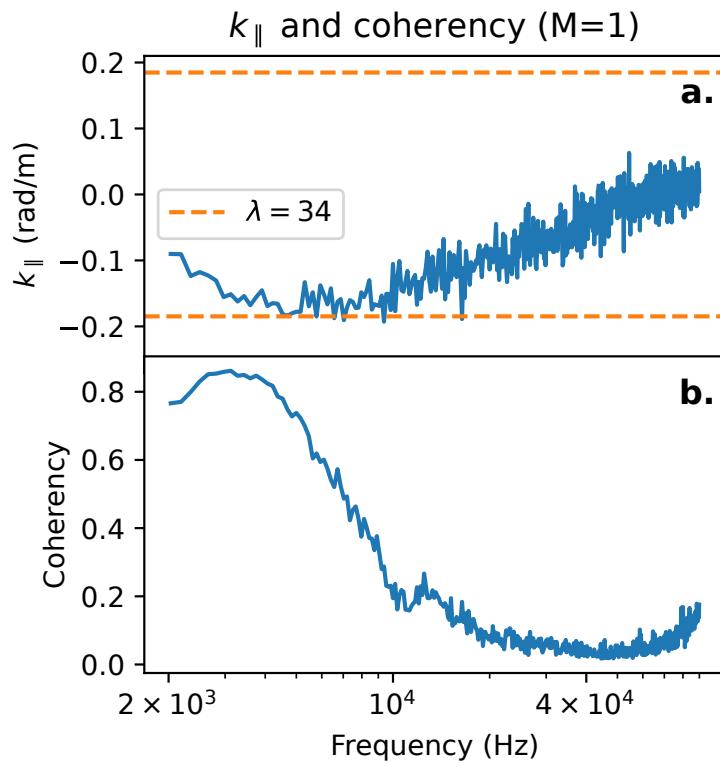


Figure 2.17: k_{\parallel} (a) and coherency γ (b) as a function of frequency. Only results from the $M = 1$ case are available, but it is clear that there are long ($\gtrsim 34$ m) wavelength modes at 3 and 12 kHz. The probes used for calculating k_{\parallel} were located at the midplane ($z=8.31$) and $z=12.14$ m, 3.83 m apart.

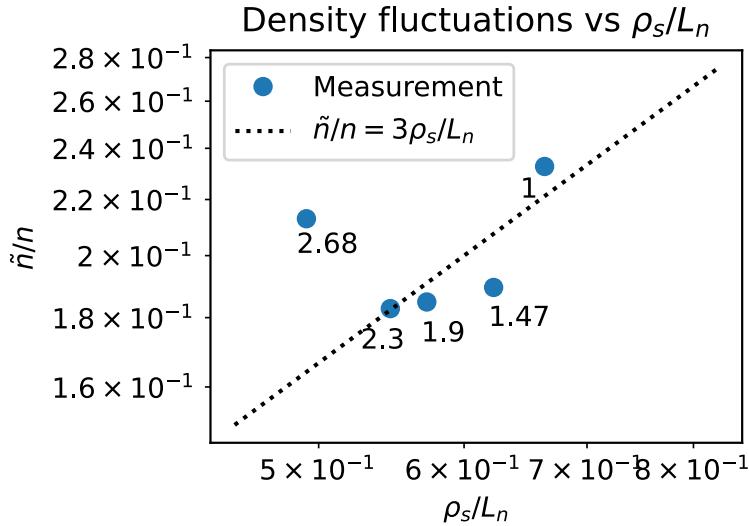


Figure 2.18: Normalized density fluctuations vs ρ_s/L_n . The measured values fall close to the $\tilde{n}/n = 3\rho_s/L_n$ line which is consistent with theory.

probes 1 and 2. The coherency is defined as $\gamma = \frac{|\langle P_{1,2} \rangle|}{\langle |P_{1,1}|^2 \rangle \langle |P_{2,2}|^2 \rangle}$ where $P_{x,y}$ is the cross-spectrum between signals x and y and the angle brackets $\langle \rangle$ denote the mean over shots. The coherency between the two Vf probes drops off with increasing frequency, with a slight bump at around 12 kHz. There are several candidates for the driving mechanism of the 3-5 kHz mode, but the 12 kHz mode is most likely a drift-Alfvén wave.

Drift wave theory[Lie85] suggests that the normalized density fluctuation level \tilde{n}/n should fall near $3-10 \rho_s/L_n$. A plot of this relation using experimental data can be seen in fig. 2.18. However, comparison of \tilde{n}/n with $1/(k_y L_n)$ show that the normalized density fluctuations are about an order of magnitude too small for the $1/(k_y L_n)$ observed which is unexpected and this conflict has not been able to be resolved in this study.

Another issue with this drift-wave interpretation of results is that the electron thermal diffusion along the field line is too high. The plasma must be collisional enough that thermal equilibrium is guaranteed (i.e., the temperature is Maxwellian), but if the collision rate is too high then thermal

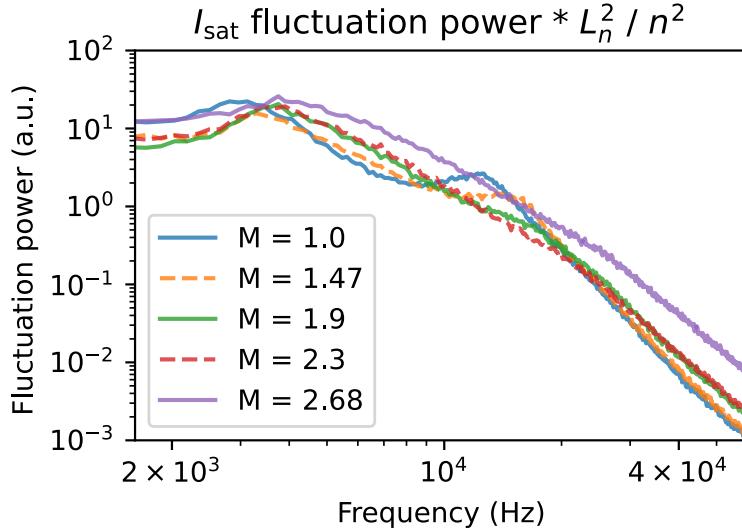


Figure 2.19: I_{sat} fluctuation power when scaled by the square of the gradient scale length and the squared density. We expect this value to be constant (assuming the same k_{\perp}).

gradients can develop along the field line[Gol95]. This condition on thermal diffusivity condition for the drift wave ω and k_z is $\omega \ll k_z^2 v_{e,th}^2 / v_{ei}$. Plugging in values from the experiments yields frequencies at least 5 times greater than mandated by the diffusivity condition and the condition is violated. This condition violation may be responsible for the odd phase shifts seen between the density and potential fluctuations.

In saturated drift wave turbulence, the normalized density fluctuation amplitude is expected to scale with the gradient scale length L_n , so the fluctuation power then scales with L_n^2 . A plot of this can be seen in fig. 2.19. This assumes the same k_y , but as mentioned earlier, that scaling and the relationship in general is not consistent with theoretical predictions for saturated drift wave turbulence.

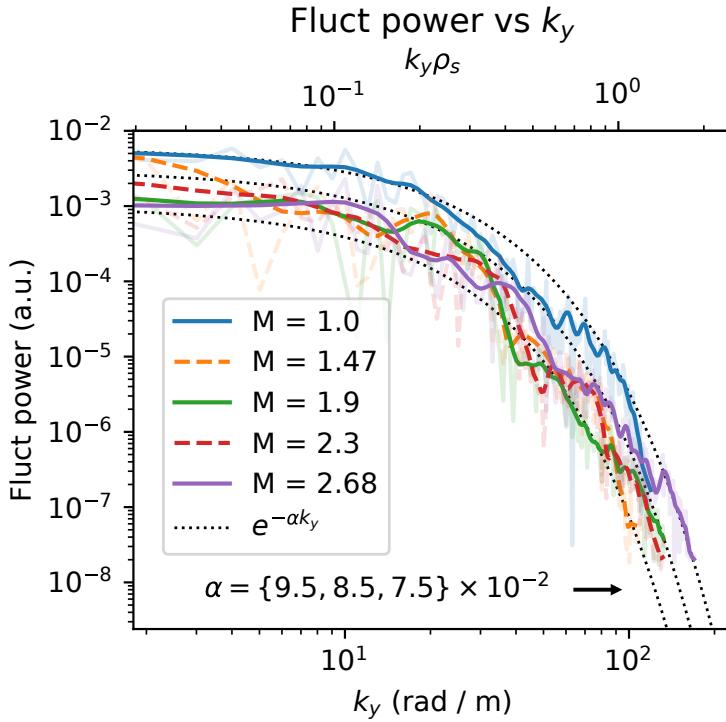


Figure 2.20: Fluctuation power summed for each k_y for frequencies up to 100 kHz, smoothed. The contribution to fluctuation power is negligible past 100 kHz. The fluctuation power decreases substantially when a mirror configuration is introduced, but no trend is seen otherwise and the k_y spectra remain exponential. Note the logarithmic scale.

2.3.5 Turbulence modification

The wavenumber-power relation in fig. 2.20 shows decreased fluctuation power when a mirror configuration is introduced. However, there is no discernible trend when the mirror ratio is increased further. The exponential nature of the curve also remains unchanged. The greatest decrease in fluctuation power occurred in low and high k_y 's, around 10 and 70 rad/m. The shape of the power- k_y curves follow an exponential distribution, and is inconsistent with a 2d drift-wave turbulent cascade (Wakatani Hasegawa k^{-3}) [WH84]. The steep dropoff in fluctuation power with k_y suggests that higher-wavenumber fluctuations do not have a significant effect on transport.

Previous simulations in a flat field [Fri13] predicted frequency and wavenumber spectra that can be fit with many power laws or exponentials, but the data presented here (figs. 2.16, 2.21, 2.20) appear to follow an exponential relationship within measurement variation.

Turbulence measurements can be directly compared to theoretical predictions and other devices, summarized by Liewer [Lie85]. For saturated drift wave turbulence, one expects the normalized fluctuation level $\tilde{n}/n \sim 1/\langle k_{\perp} \rangle L_n$, where k_{\perp} is some typical wavenumber. The power-weighted k_y (calculated from fig. 2.20) was approximately 15 rad/m, which is an order of magnitude too small to satisfy this relationship. \tilde{n}/n scaling with ρ_s/L_n , however, is roughly consistent with drift wave turbulence level saturation: the latter is ≈ 3 times larger. These comparisons suggest that the large, low frequency fluctuations (~ 3 kHz, which had even smaller k_y) may have a drift wave turbulence component but are dominantly driven by other instabilities. No trend is seen in ρ_s/L_n and $1/k_y L_n$ when mirror ratio was varied.

Core fluctuations appear to decrease dramatically as seen in the I_{sat} fluctuation power (fig. 2.8). The I_{sat} decorrelation time increases from ~ 0.7 ms for $M = 1$ to ~ 2.5 ms for $M = 2.68$. At $x = x_{PF}$, decorrelation times for all mirror ratios remained at 0.2 ms.

2.3.6 Magnetic fluctuations

The perpendicular magnetic fluctuation (B_{\perp}) component of the drift-Alfvén wave can be seen in fig. 2.21. These B_{\perp} fluctuations are spatially and spectrally coincident with the electrostatic fluctuations (fig. 2.16). Drift-Alfvén waves have been studied in the LAPD in the past [MM97, VG06]; strong coupling is observed for $\beta_e > m_e/m_i$ which is satisfied in this study. The Alfvén speed $\omega/k_{\parallel} = v_A = B/\sqrt{4\pi n M}$ (given $\omega \ll \Omega_{ci}$) when averaged over the entire column ranges from ~ 450 to ~ 1600 km/s. A k_{\parallel} corresponding to a wavelength $\lambda = 34$ m roughly falls within the bound established by the kinetic and inertial Alfvén wave dispersion relations at the frequency peaks observed at $x \sim x_{PF}$ seen in fig. 2.21. The lengthening of field lines caused by curvature accounts for at most 10% of the change in frequency.

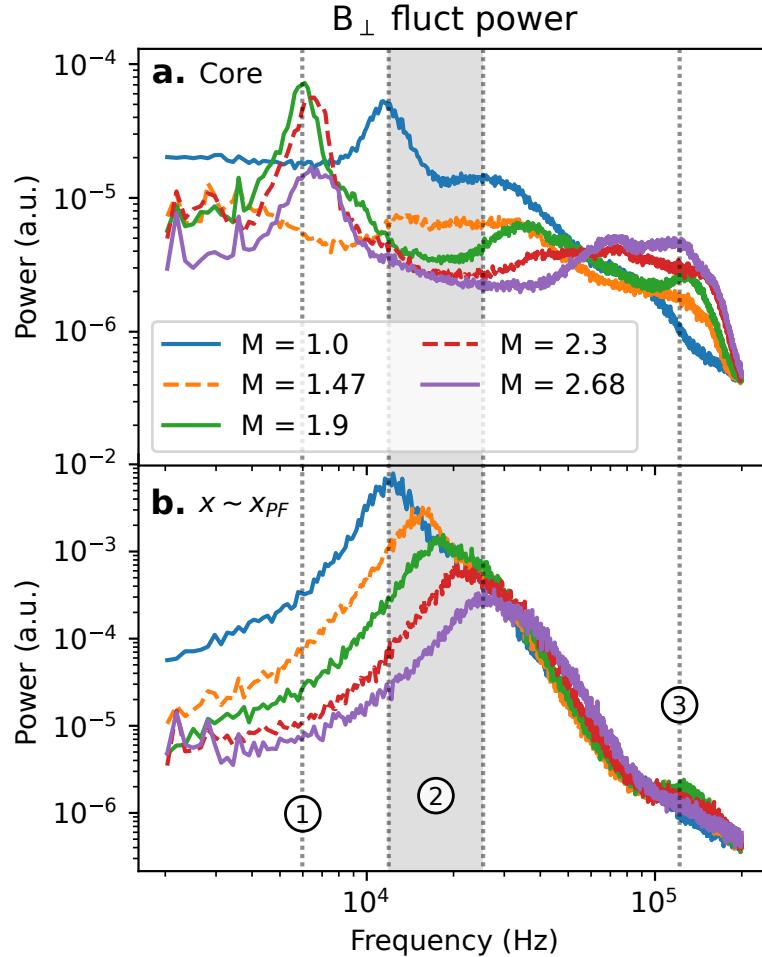


Figure 2.21: B_{\perp} fluctuation power averaged at the core from 0 to 3 cm (a) and around the peak fluctuation point ($x \sim x_{PF}$) (b). Fluctuation power decreases across the board with mirror ratio except for core frequencies close to Ω_{ci} . Peaks around 10 – 30 kHz at x_{PF} are consistent (region 2) with drift-Alfvén waves and the near-cyclotron frequency features in the core may be resonating Alfvén waves created by the magnetic mirror. Frequencies below 2 kHz and dominated by instrumentation noise and thus excluded.

The spatial extent of the B_{\perp} features identified in fig. 2.21 are plotted in fig. 2.22. Feature 1 at ≈ 6 kHz shows increased fluctuation amplitudes at $x = 0$ for mirror ratios 1.9 and above, but for $M = 1$ and $M = 1.47$ there is no increase in fluctuation power. A similar feature, but at a much smaller level, is observed in I_{sat} fluctuation power in the core as well. This core feature may be caused by the hole in the core seen in the I_{sat} profile (fig. 2.4) driving low-amplitude waves or instabilities. Feature 2 in fig. 2.22 is the magnetic component of the drift-Alfvén wave. The fluctuation power peaks at the gradient region and corresponds with the peak in density fluctuations (fig. 2.8).

Feature 3 is particularly interesting because this the only fluctuating quantity to *increase* with mirror ratio, seen in fig. 2.23. This feature may be broad evanescent Alfvénic fluctuations from the plasma source. These fluctuations have been observed in the LAPD in the source region alongside an Alfvén wave maser [MMC05]. Note that the Alfvén maser cannot enter the mirror cell at mirror ratios greater than 1.75 because the Alfvén maser resonates at $0.57 f_{ci}$ but the midplane is always at or near 500G.

The sub-2 kHz modes in B_{\perp} and its harmonics are nearly constant in power across the entire plasma; these features are likely perturbations from the magnet power supplies and thus ignored. The lack of radial, azimuthal, and axial structure in these magnetic signals below 2 kHz and narrow bandwidth indicate a non-plasma origin. Significant radial and azimuthal structure in B_{\perp} fluctuation power starts to appear in frequencies larger than 4 kHz.

The drift-Alfvénic nature of the 12 kHz Bdot feature is confirmed by changing the flat field from 500G to 400G: the feature shifts down in frequency from 12 to 10 kHz seen in fig. 2.24. From the drift wave and Alfvén wave dispersion relations the frequency is expected to be $400G / 500G = 0.8$ of the original, which is approximately what is observed. The k_y of the drift-Alfvén wave also has an effect and may be responsible for a $10 \text{ kHz} / 12 \text{ kHz} = 0.83$ factor instead.

There may be some sort of resonator made by the mirror cell and its interaction with Alfvén waves. In fig. 2.25, the behavior of the B_{\perp} spectrum in the core changes dramatically between 1 and 10 kHz in the short mirror when compared with the medium and longer mirrors. It's unlikely

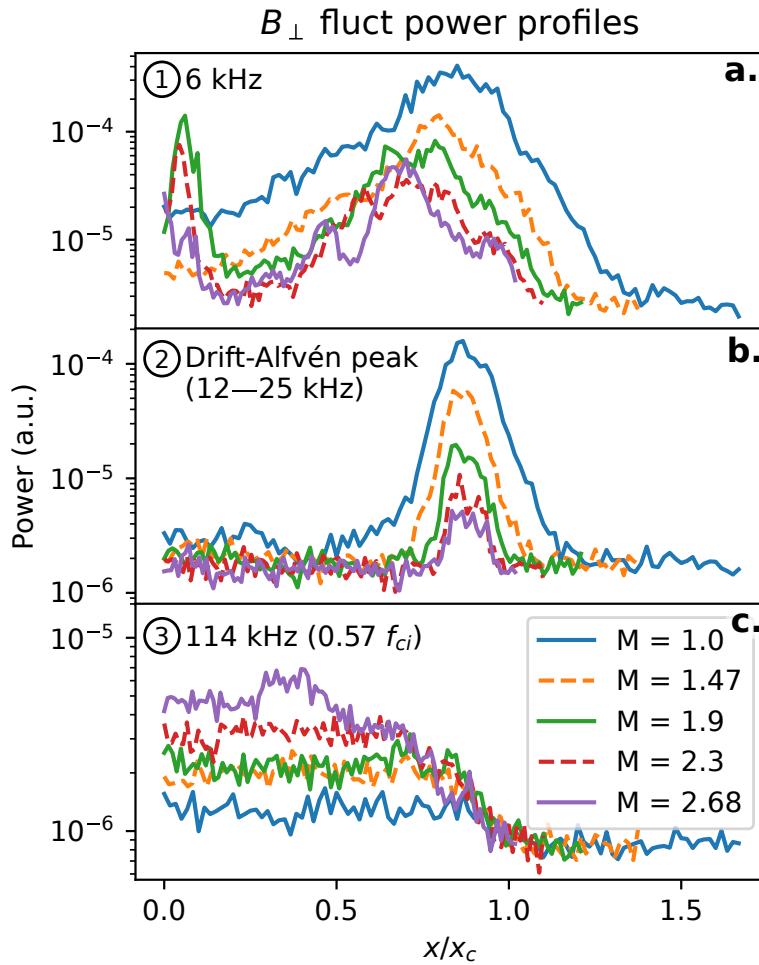


Figure 2.22: B_{\perp} fluctuation power profiles for the three regions shown in fig. 2.21: region 1 (6 kHz) (a), region 2 where frequencies are taken from the peaks of the drift-Alfvén waves for each mirror ratio (b), and region 3 (114 kHz) (c).

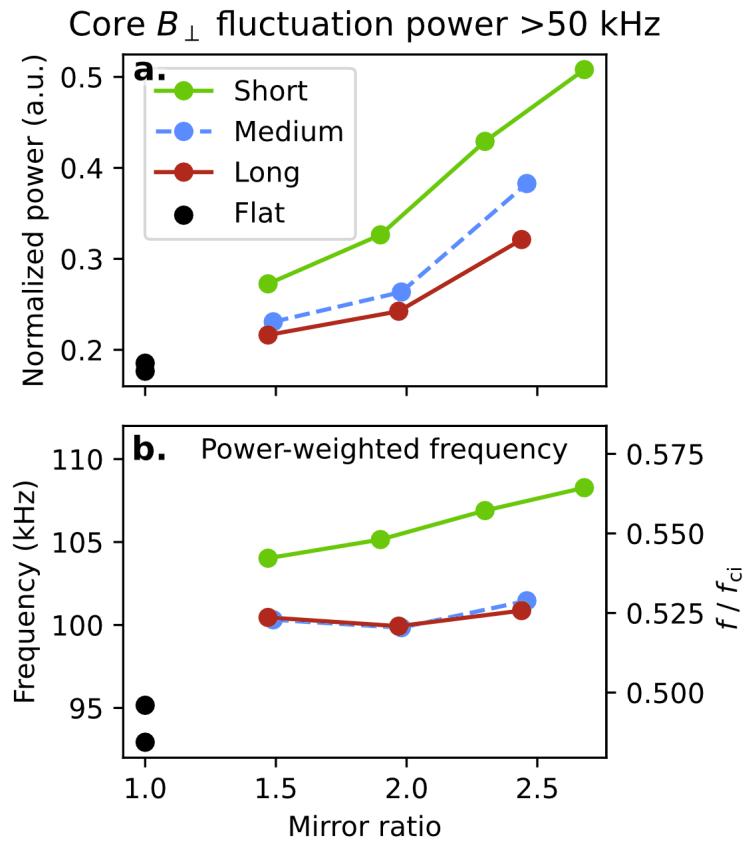


Figure 2.23: Summed fluctuation power of B_{\perp} in the core ($x/x_c \leq 0.3$) as a function of mirror length and ratio. Top (a): the fluctuation power is normalized by the sum of the full-spectrum summed power. Bottom(b): the frequency of the power distribution > 50 kHz weighted by the fluctuation power.

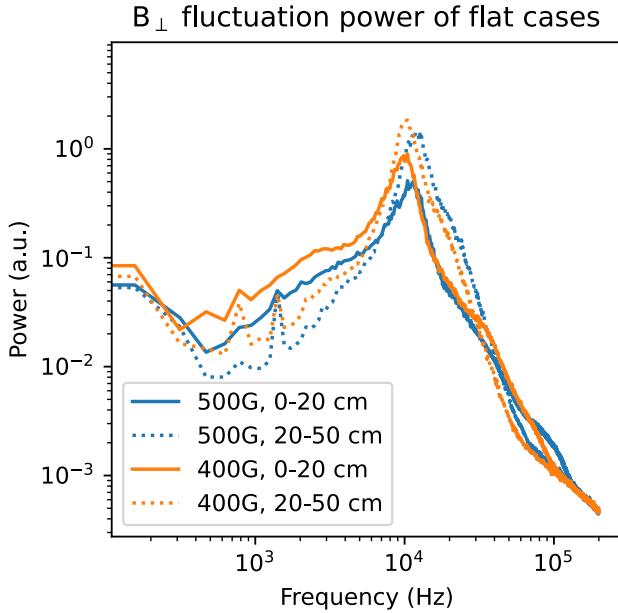


Figure 2.24: B_{\perp} flat field for 500G and 400G flat fields. The frequency of the identified drift-Alfvén wave at 12 kHz drops when the field is lowered, as expected.

that this is an Alfvénic fluctuation because the wavelength is an order of magnitude too large to fit in the machine.

For completeness, B_z fluctuation measurements are seen in fig. 2.26. The peaks in the 10 kHz region are likely crosstalk or slight coil misalignment of the probe and are picking up B_{\perp} fluctuations. The profile low frequency B_z fluctuations can be seen

The low frequency fluctuations in the B_{\perp} spectra may seem important but plotting the spectra as a function of position (fig. 2.27) clearly shows the harmonics of the signal and the narrow bandwidth of them. This spectral feature is present regardless of mirror ratio, but changes in magnitude in approximate proportion with the field, i.e., the magnet power supply current. This power supply-induced field fluctuation can easily be seen in the ≈ 625 Hz mode in B_z , seen in fig. 2.28. The fluctuation power is largely constant across the entire plasma column, with the fluctuation power increasing with increased mirror fields. The taper of the fluctuation power at

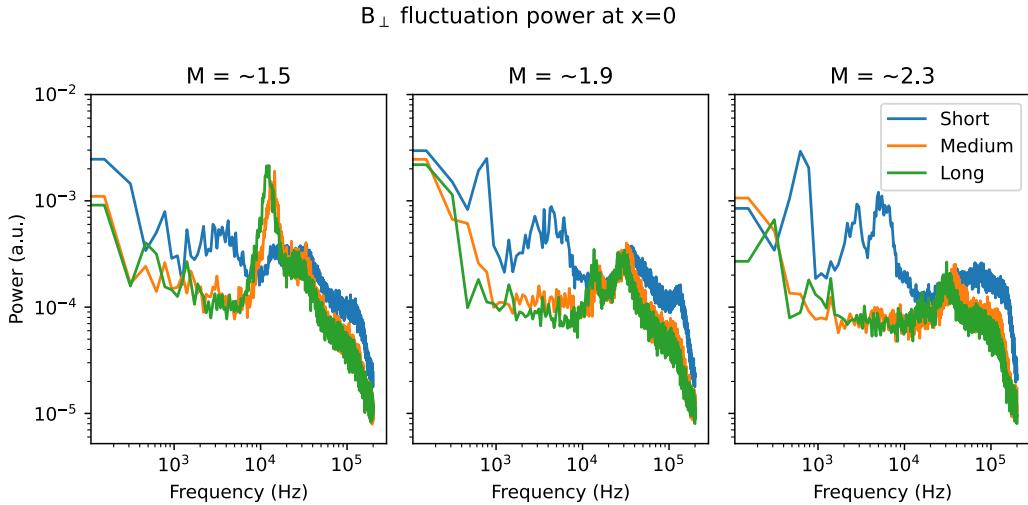


Figure 2.25: B_{\perp} at $x=0$ for different mirror lengths. The origin of fluctuations between 1 and 10 kHz is unknown.

the edge could be caused by the background field vector no longer pointing in the z direction as the probe approaches the magnet coil. In general, the probe valves are not centered between the magnet coils, leading to rotation of the field vector as the probe is pulled out.

2.4 2d Structure

The perpendicular magnetic field structure is measured by collecting x-y planar bdot ($dB_{\{x,y,z\}}/dt$) data alongside a stationary, axially separated I_{sat} reference probe (DR2). This probe provides a phase reference for the magnetic field fluctuations, allowing a 2d map of relative phase to be constructed over many shots. Only the region around x_{PF} was measured because of constraints on probe movement. The amplitude and phases for each magnetic field component are then used to reconstruct the local magnetic fluctuation vector \mathbf{B} . The axial current density structure, j_z , can be derived from this vector field. \mathbf{B} and the corresponding j_z for the flat-field ($M = 1$) case can be seen in fig. 2.29. Two main current channels can be seen with the magnetic fields circulating around them. This structure quickly decoheres in time as expected in a turbulent plasma. At higher

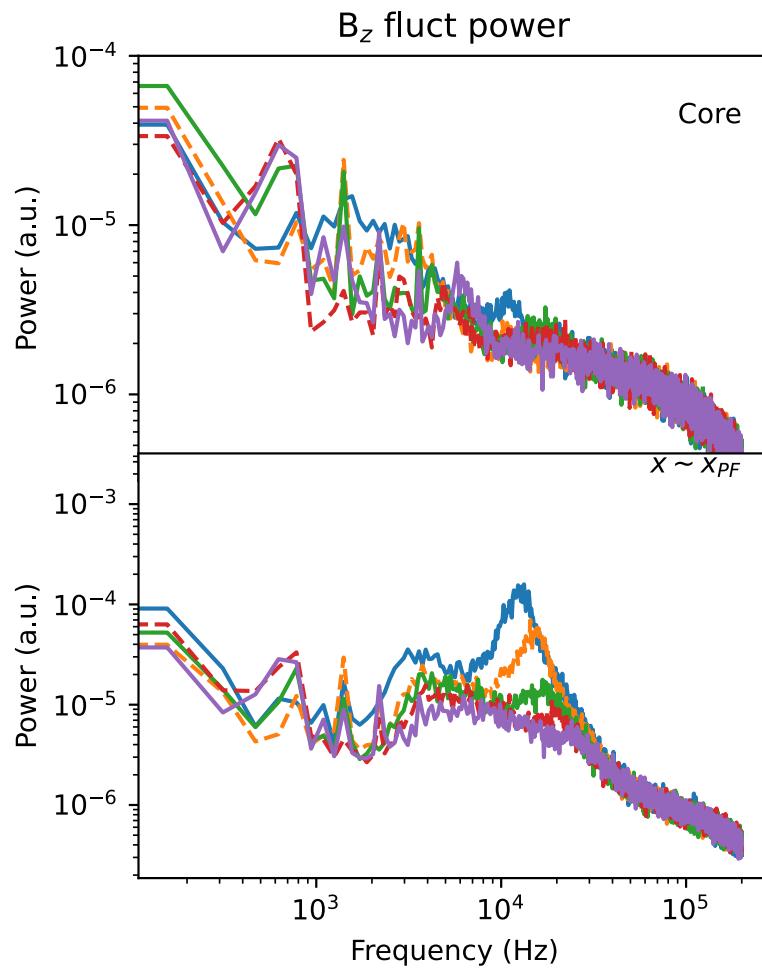


Figure 2.26: B_z fluctuations in the core and x_{PF} . Aside from picking up some B_\perp signal, the spectra are largely featureless.

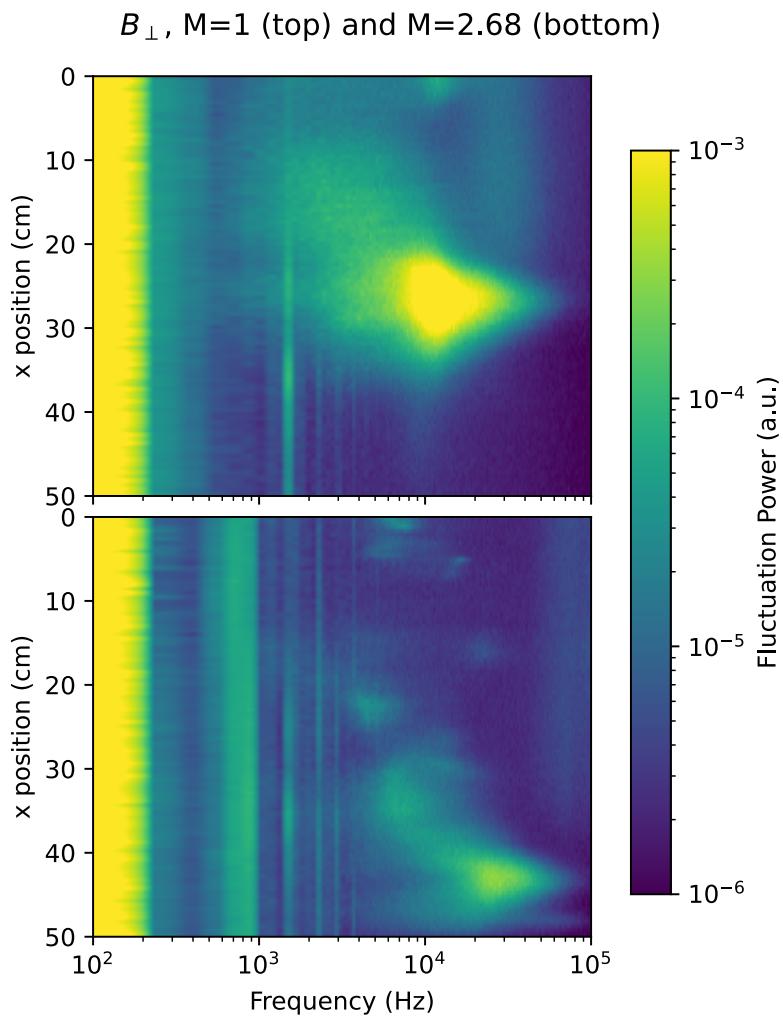


Figure 2.27: B_{\perp} fluctuation power for mirror ratios of 1 and 2.68. Lower frequencies are shown and the colorbar clipped to show detail in what appears to be power supply fluctuations.

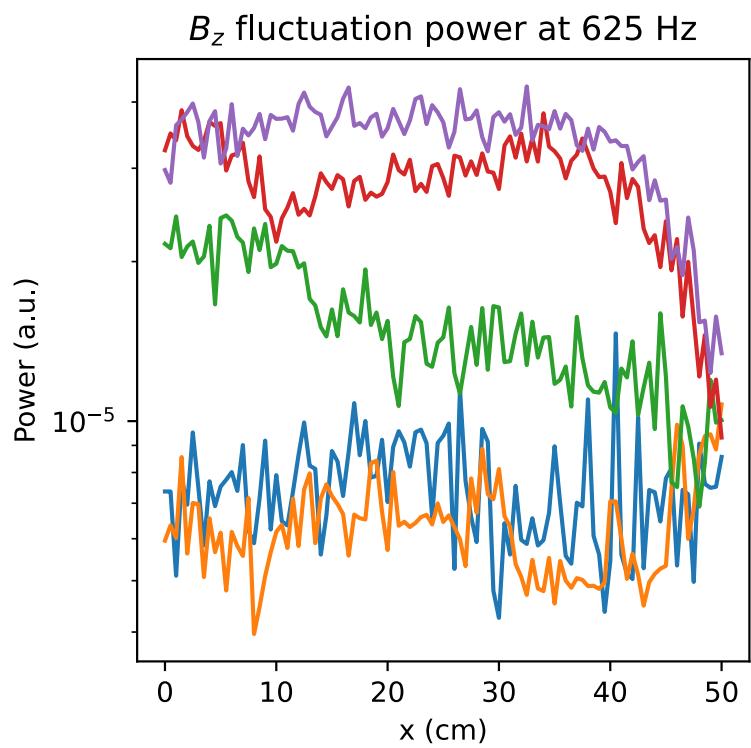


Figure 2.28: B_z fluctuation power profiles for all mirror ratios at 625 Hz.

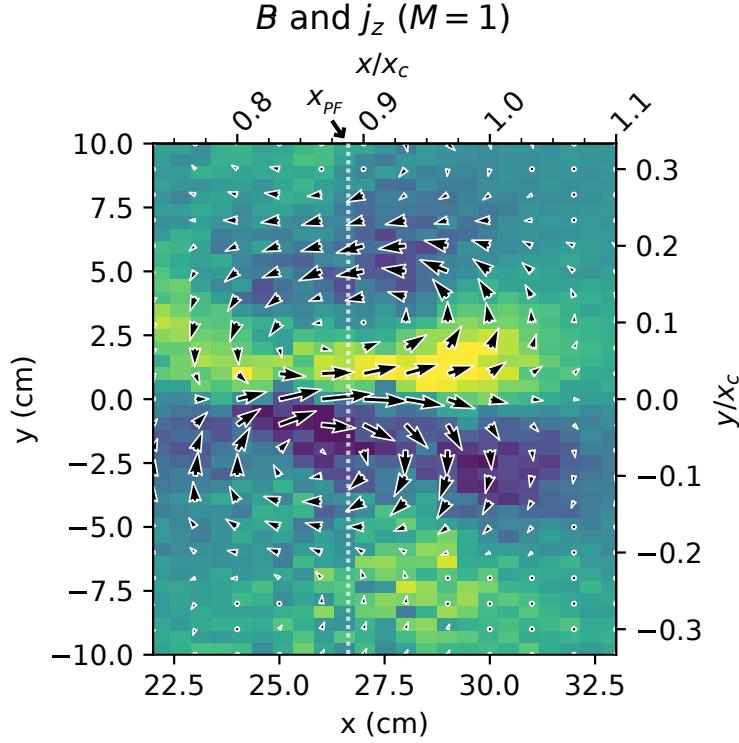


Figure 2.29: Perpendicular magnetic field and the derived current density for the flat-field ($M = 1$) case using a Bdot probe with an axially-separated I_{sat} reference (DR2). The x-y plane was centered near x_{PF} .

mirror ratios, the field magnitude and corresponding current density decrease (which was also seen in DR1: fig. 2.21). Similar structure is seen in the mirror configurations; the $M = 1.9$ and $M = 2.68$ cases can be seen in fig. 2.30.

Using two, axially-separated, correlated I_{sat} measurements (DR2), with one collecting x-y planar data, the azimuthal mode number m (radially integrated) was calculated. Higher-frequency and higher- m features are seen with increasing mirror ratio (fig. 2.31). The increased frequencies may be caused by a change in Doppler shift by the $\mathbf{E} \times \mathbf{B}$ flow. This higher- m trend suggests that azimuthal structures do not scale with increased plasma radius but instead remain roughly the same size. The limited planar probe movement caused an increase in the lower bound on m in higher

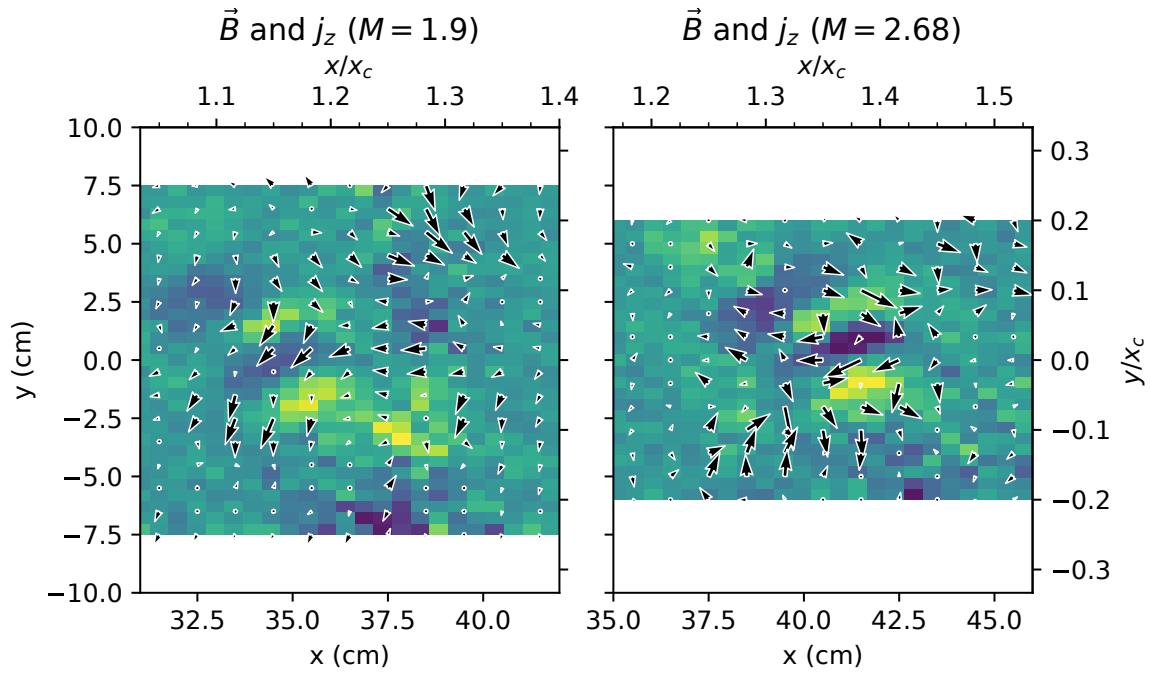


Figure 2.30: Perpendicular magnetic field and the derived current density for the $M = 1.9$ and $M = 2.68$ cases computed in the same manner as fig. 2.29. The x-y planes were centered near x_{PF} , and the view size was kept constant across the plots. The structure is much less obvious in the mirror cases, but all exhibit the expected Alfvèn wave pattern

mirror ratios. At mirror ratios 1.47 and higher, the lower frequency component (< 10 kHz) appears to decrease significantly in amplitude. Calculating k_{\perp} from m evaluated at $x \sim x_c$ yields similar k_y values as the two-tip technique (fig. 2.32). The average k_y for a given frequency can be calculated using two Vf tips on the same probe by calculating the phase difference and dividing by the spatial separation of 5 mm: $k_y = \phi_{vf1, vf2}/\Delta y$ [BKP82]. The maximum $|k_y|$ measurable before aliasing is $\pi/\Delta y \approx 628$ rad/m. As seen in fig. 2.32, the k_y spectrum remains similar across mirror ratios, but the wavenumber extends further into higher frequencies with increasing mirror ratio. These azimuthal mode numbers and gradient scale lengths are consistent with linear simulations using the 3d fluid code BOUT [PUC10] in the flat, unbiased case.

2.5 Discussion

2.5.1 Lack of mirror-driven instabilities

No evidence is seen for mirror-driven instabilities curvature, loss-cone, or otherwise. Given the LAPD parameters in this study (tables 2.1 and 2.2), the collision frequencies are sufficiently high such that the mirror is in the gas-dynamic regime: losses out of the mirror throat are governed by gas-dynamic equations rather than free streaming through the loss cone. To be in the gas-dynamic regime, the mirror length must exceed the mean free path of the ions [IP13]:

$$L > \lambda_{ii} \ln M/M \quad (2.2)$$

where L is the mirror length, λ_{ii} is the ion mean free path, and M is the mirror ratio. These collisions populate the loss cone and maintain a (cold) Maxwellian distribution, eliminating the possibility of loss-cone-, ion-driven instabilities like the AIC [CS82] or DCLC [Sim76, Kan79] instabilities that have been observed in other (historic) devices.

The paraxial, approximate interchange growth rate is [Pos87, RBC11]

$$\Gamma_0 = \frac{c_s}{\sqrt{L_M L_P}} \quad (2.3)$$

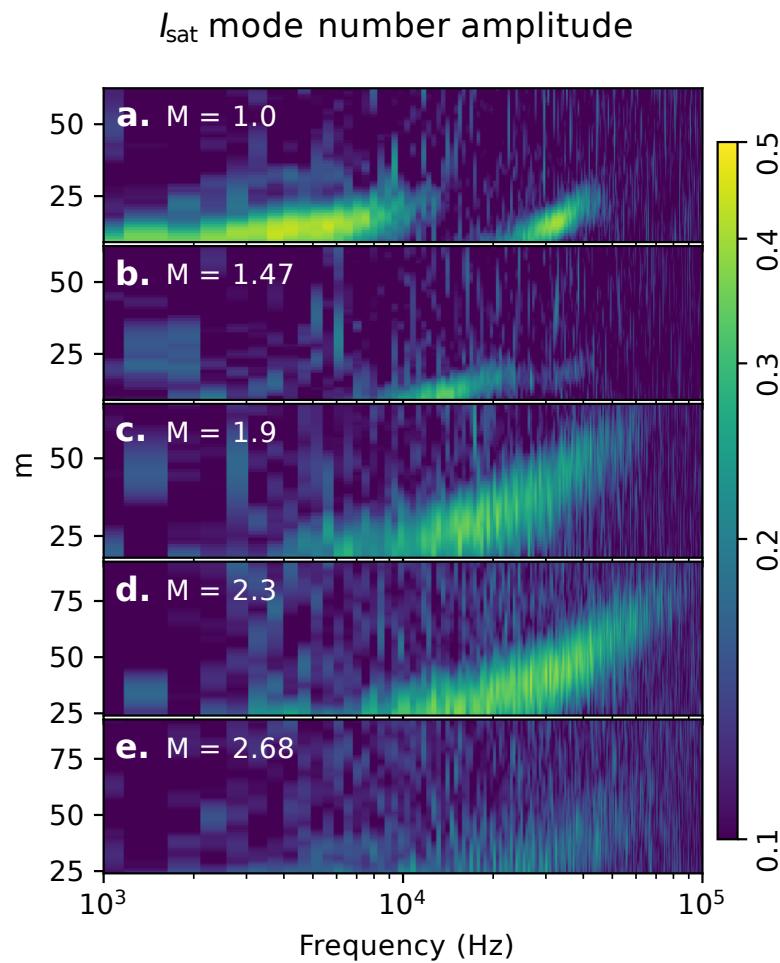


Figure 2.31: Azimuthal mode number m amplitudes calculated from two axially-separated, correlated, I_{sat} probes. Increasing mirror ratio (a to e) leads to increased m at higher frequencies.
(DR2)

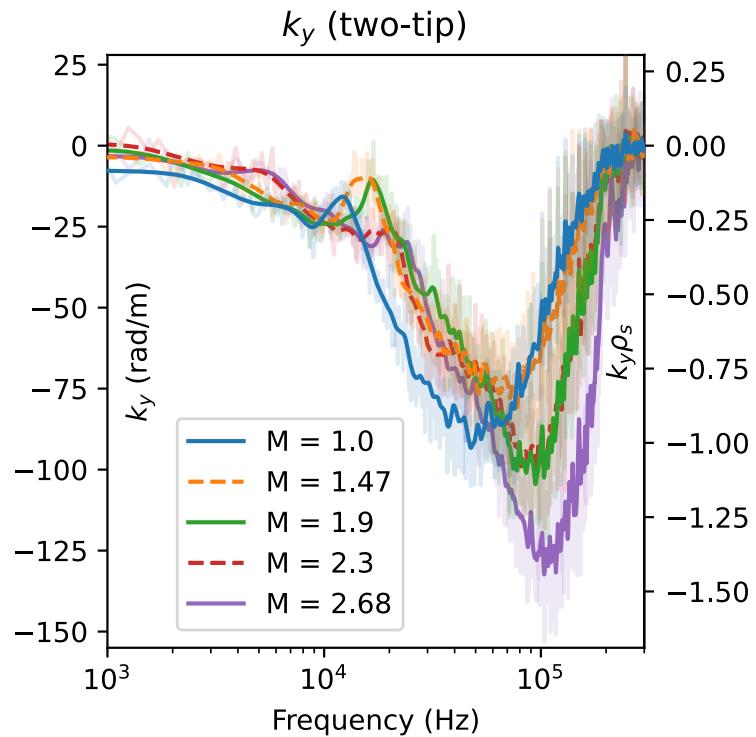


Figure 2.32: k_y averaged about x_{PF} and smoothed for each mirror ratio calculated using two vertically-separated Vf tips on the same probe. Little change is seen in k_y at lower frequencies but higher frequencies tend towards larger k_y at higher mirror ratios.

which yields $\Gamma_0 \approx 1.2$ kHz using $L_M \approx 7$ m and $L_P = 17$ m. c_s is used instead of \bar{v}_i because $T_i \ll T_e$ and mirror length L is split to distinguish between the contributions of the plasma length and mirror length to inertia and to curvature drive, respectively. Interchange is not visible in-part because the aspect ratio of these mirrors is quite large, limiting the growth rate of interchange. The length of the mirror (3.5 m), radius of curvature (6-7 m), and plasma column (17 m) are much larger than the radius of the plasma (0.5 m maximum), so the plasma inertia is large relative to the instability drivers. Line-tying to the cathode may further lower the growth rate. The hot cathode used for plasma formation could function as a thermionic endplate that can supply current to short out the flute-like interchange perturbations. Line-tying has been seen in flux rope experiments on the LAPD using a hotter, denser source [VGP11], also in other devices [FKR79], and is why interchange was not seen in the earliest mirror machines [Pos87]. Note that the plasma terminates on the cathode or end plates before the magnetic field flares out, so there is no contribution to stability from an expander tank as seen in other GDTs [RBC11, IP13]. Finite Larmor radius (FLR) effects may provide a stabilizing effect for larger azimuthal mode numbers. At the highest mirror ratio, assuming a plasma radius of $a_0 = \sqrt{2.68} * x_{PF} = 43$ cm, the FLR stability criterion $\frac{m}{2} \frac{\rho_i L}{a_0^2} > 1$ [RBC11] suggests a stabilizing effect may be present for azimuthal mode numbers $m > 4$.

If the curvature-induced interchange instability were observable, then introducing a mirror configuration would lead to new features in I_{sat} and B_{dot} fluctuations. In particular, low-frequency mode(s) – likely less than 10 kHz given the low m-number and plasma rotation rates – would be observed growing from the pressure gradient region. For onset of the interchange instability, the mirror curvature or plasma pressure would need to be increased but the precise conditions required for this onset are not yet known for the LAPD.

Interchange could also be at least partially stabilized by the continuous production of electrons in the core that are electrostatically trapped by the ambipolar potential [GH71]. The intuition behind this stabilization mechanism is as follows: electrons are continuously produced via ionization of neutrals, and any change in the local potential will cause more or fewer electrons to be lost out the ends of the device along that field line, counteracting the potential change. This stabilization

mechanism has been experimentally demonstrated to completely suppress interchange when the ambipolar potential $\Phi \gtrsim 6T_e$ [KHS87].

The $\mathbf{E} \times \mathbf{B}$ shear flow present (fig. 2.7) may also make a contribution to the stabilization of interchange [RBC11, BLZ03, BBS07, BBC10]. The estimated shearing rate is between 3 and 10 kHz, which is greater than the estimated ≈ 1.2 kHz growth rate of the interchange mode.

2.5.2 Instabilities driving LAPD turbulence

Rotational interchange can be significant driver of the broadband turbulence spectrum in the LAPD, particularly when a biased limiter is installed. This observation has been confirmed by both linear simulations [PUC10] and biasing experiments [Sch13].

This rotational interchange mode has the following attributes, as summarized by [Jas72]: flute-like ($k_{\parallel} = 0$), $|e\tilde{\phi}/T_e|/|\tilde{n}/n| \gtrsim 1$, radial potential phase variation 45 to 90°, maximum possible $|e\tilde{\phi}/T_e| < 1$. All of these attributes are seen for the lower frequency (3 kHz) mode. The Vf radial phase variation when $M > 1$ is not clearly seen because the coherency is dramatically reduced along the field line. The rotational interchange mode could couple with the drift wave at $k_{\parallel} = \pi/L \sim 0.37$ rad/m ($n = 0.5$), which has been observed in the past [Sch13] and likely present here. Estimates of shearing rate from the $\mathbf{E} \times \mathbf{B}$ flow velocity profile (fig. 2.7), calculated fluctuation ratios, and radial phase shift variation suggest that Kelvin-Helmholtz-driven turbulence is not significant, if present at all. Historically, biasing a limiter has been required to clearly observe the Kelvin-Helmholtz instability [HPC05, SCR12, Sch13].

Low frequency density fluctuations may also be driven by a flute-like conducting-wall temperature-gradient instability which only requires an electron temperature gradient to grow (even with straight field lines) [BRT91]. Simulations of turbulence in the LAPD suggest the possible presence of these conducting wall modes (CWM) which have the highest growth rate for $m \leq 20$ [FCU13]. This lower- m mode could be responsible for the peak around 3 kHz in the $M = 1 I_{\text{sat}}$ fluctuation (fig. 2.16) and azimuthal mode numbers (fig. 2.31) and for the low-frequency low- k_{\parallel} or flute-like

behavior (fig. 2.17). This CWM may also be responsible for flatter electron temperature profiles seen in previous studies [PMC22, Sch13] (fig. 2.6).

These linearly unstable modes may be outgrown by a rapidly-growing nonlinear instability that couples to drift-like modes as suggested by simulations [FCU13]. This nonlinear instability is driven by the density gradient at an axial modenumber of $n = 0$ and nonlinearly transfers energy to $n \neq 0$ fluctuations.

The conducting wall mode and nonlinear instability may be present in these mirror experiments but the spectra are adequately explained by linearly unstable modes. Precise identification of these modes requires further study; neither of these instabilities have been directly observed in the LAPD.

2.5.3 Causes of particle flux reduction

The reduction in particle flux explained by a reduction in density fluctuations likely caused by a increased gradient scale length $L_n = \frac{n}{\nabla n}$ (fig. 2.33), decreasing the linear drift wave growth rate and saturation level seen in sec. 2.3.5. This gradient length reduction may also reduce the growth rate of the rotational interchange instability, which may be the dominate driver for the low-frequency large-amplitude density fluctuations. The influence of this density fluctuation reduction appears reduced at higher mirror ratios past $M = 1.9$, where the wavenumber and phase angle appear to decrease in magnitude. The plot showing this breakdown in particle flux can be seen in fig. 2.34. The changes in I_{sat} fluctuation power is the most obvious driver, but the $I_{\text{sat}} - V_f$ phase difference, coherency and wavenumber also seem to have an effect. The V_f fluctuation power remains largely consistent across the different mirror ratios. Note that this particle flux appears somewhat different because this is using the uncalibrated I_{sat} values and the flux is not scaled by solid angle. This flux also does not use temperature-compensated I_{sat} measurements.

The decorrelation time of I_{sat} time series data is around 0.15 ms at x_{PF} . An estimate of the $\mathbf{E} \times \mathbf{B}$ flow shear from fig. 2.7 (DR2) yields a shearing time between 0.1 and 0.3 ms at x_{PF} . These

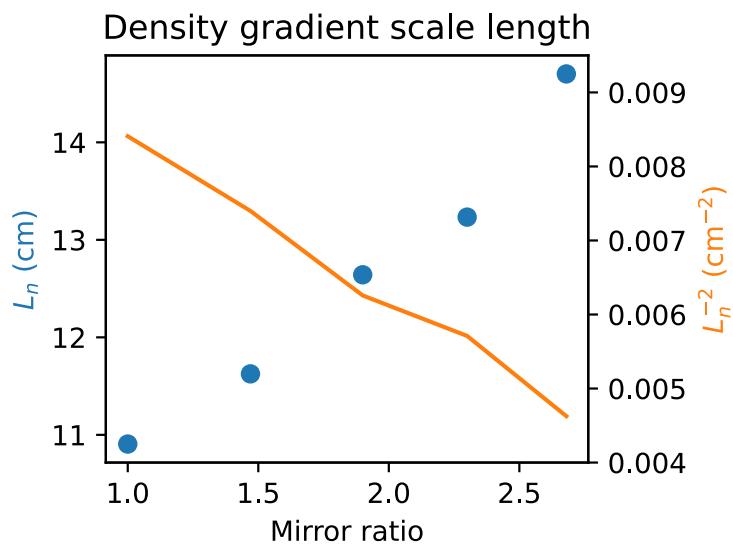


Figure 2.33: Gradient scale length L_n and the associated term in the drift wave growth rate L_n^{-2} . This scale length was calculated over a 3 cm region around x_{PF} (peak fluctuation region) at the mid-plane. Increasing the mirror ratio increases the gradient scale length, which suggests weakening of the underlying instability driver.

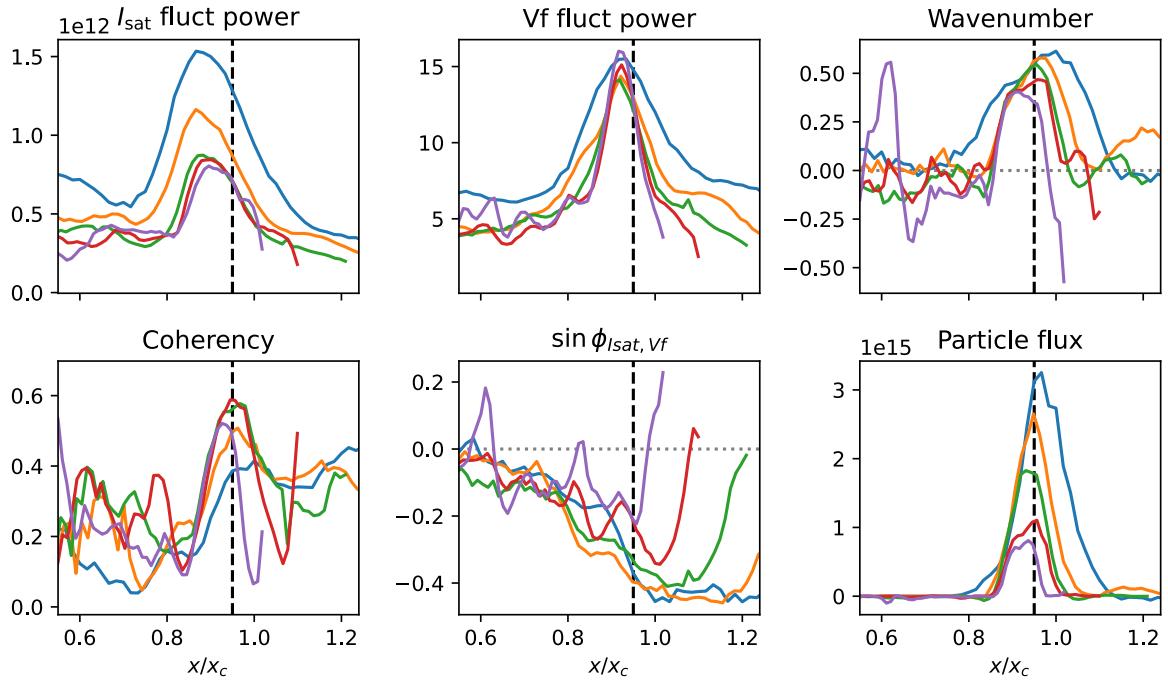


Figure 2.34: The particle flux broken down into the components used to calculate it. The dashed black line is simply a visual reference near the peak particle flux at $x/x_c = 0.95$. The I_{sat} fluctuation power appears to be the largest driver in changes in particle flux. The colored lines correspond to mirror ratio as seen in earlier plots.

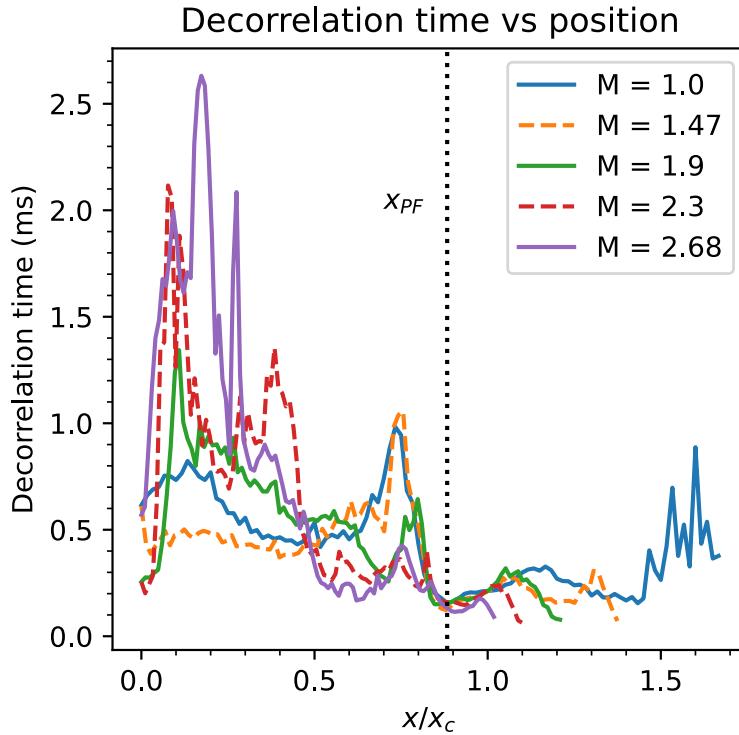


Figure 2.35: Decorrelation time from I_{sat} time series data for different mirror ratios. All of the mirror ratios have a minimum decorrelation time at x_{PF} and much longer times (slower rate) in the core.

times suggest that spontaneous flow shear may be important for suppressing turbulence, as seen in other studies [SCR13, CYK05], at all mirror ratios. However, no clear trend in shearing strength is seen with mirror ratio.

The decorrelation time of a signal is calculated by taking the autocorrelation of a signal – I_{sat} in this case – and finding the full-width half-max of the envelope using a Hilbert transform. This decorrelation time can be seen in fig. 2.35. The decorrelation is minimized at x_{PF} and maximized in the core, further confirming the turbulent nature of the fluctuations at x_{PF} .

The estimated shearing rate from DR2 can be seen in fig. 2.36. The rate is plotted instead of time because of the singularity when the flow reverses. At around x_{PF} ($x/x_c \approx 0.87$), the shearing

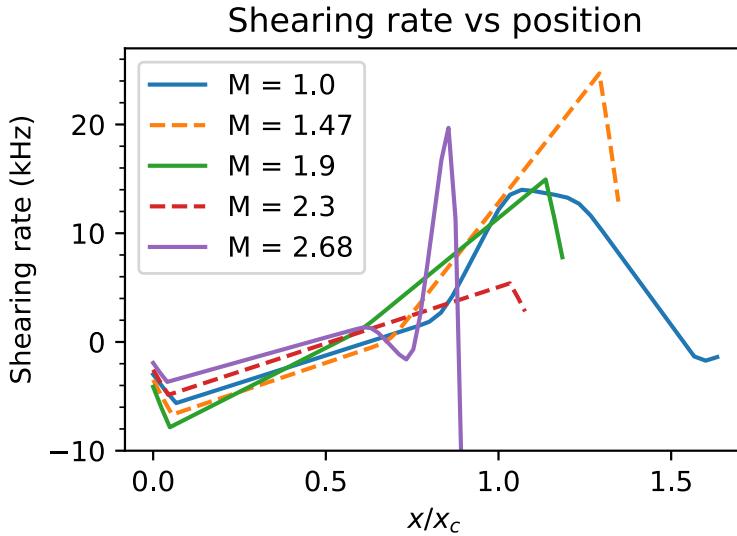


Figure 2.36: ExB shearing rate based on smoothed plasma potential profiles. This shearing rate is comparable to the decorrelation rate at x_{PF} , seen in fig. 2.35.

rate is around 2 to 8 kHz meaning the shearing time is around 0.5 to 0.125 ms. This is fairly close to the decorrelation time from the I_{sat} time series measurements (fig. 2.35). These similar times/rates suggests that ExB shearing may set the limit on cross-field transport.

It is important to note that the electron thermal diffusion time along the field line is very long compared to the frequency of the drift wave ($\omega \gtrsim k_{\parallel} \bar{v}_e^2 / v_{ei}$) [Gol95] so the electron temperature along the field line may not be constant on the drift wave timescale. This factor is not taken into account in this analysis but may have substantial impact on interpretations of the measured phase shift.

2.5.4 Differences between DR1 and DR2

Directly applying signals between these two dataruns is not quite appropriate because the profiles/plasmas changed appreciably. These changes could have been caused by differences in cathode temperature, emissivity, or other properties. The discharge power for DR2 was roughly 10%

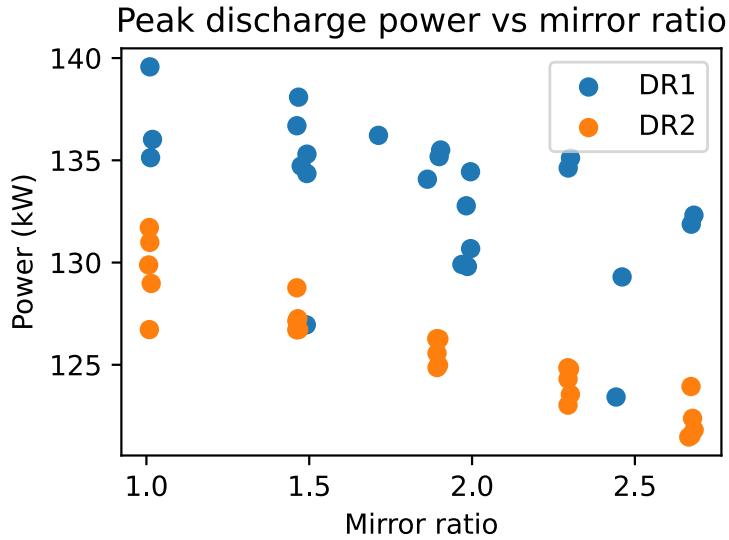


Figure 2.37: Discharge power vs mirror ratio. DR2 had roughly 5% lower discharge power than in DR2 for unknown reasons.

smaller than what was seen in DR2 seen in fig. 2.37. Since the discharge voltages were similar (DR1: 62.5 vs DR2: 60.5) we expect to see less dense plasmas in DR2.

Changes in the I_{sat} profiles between the two dataruns (and between two separate measurements in DR1) can be seen in fig. 2.38. Interestingly, there is some difference in the profiles *within the same datarun* which could be caused by probe shadowing. Probe shadowing effects should be less important in mirrors because the probe closest to the cathode magnetically maps to a region further outside than the probes in the mirror cell. This difference in density can also be seen in the line-integrated density from the 56 GHz interferometer (port 23): fig. 2.39. These differences in density could also be caused by different hydrogen and helium pressures in the runs. Helium pressure was roughly the same for both dataruns (6×10^{-6} to 3×10^{-6} for DR1, 6×10^{-6} to 2×10^{-5} for DR2), but the hydrogen pressure was an order of magnitude higher for the DR2, on the order 7×10^{-6} instead of 1×10^{-7} for DR1. These differences in pressures could have had an effect on plasma formation and transport, thus affecting profiles. Hydrogen fraction is known to have an effect on breakdown characteristics in the newer Lanthanum-hexaboride (LaB6) cathode.

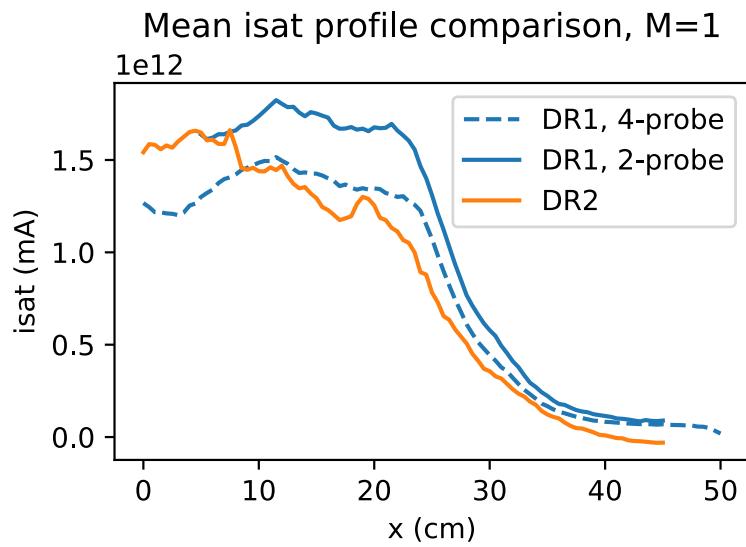


Figure 2.38: I_{sat} profiles ($M=1$), DR1 vs DR2 in the mirror cell. DR2 has reduced I_{sat} across the entire profile. Identical machine settings in DR1 a few days apart are also different.

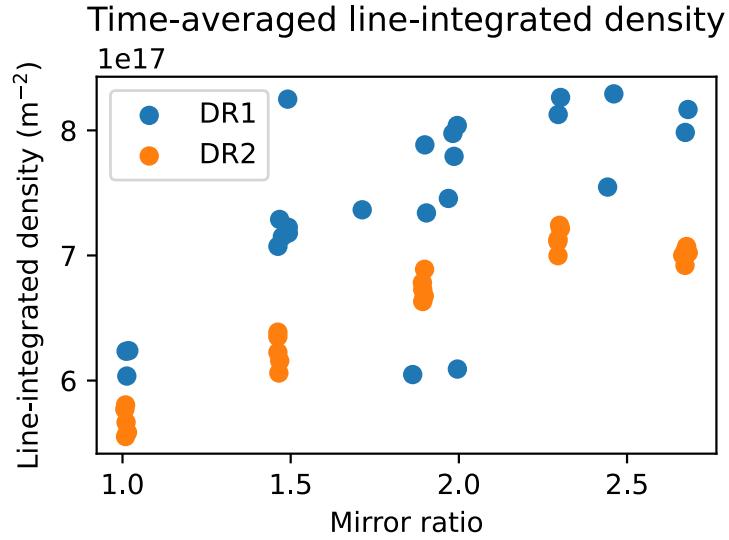


Figure 2.39: Line integrated density from the 56 GHz, DR1 vs DR2. On average, DR2 has a lower density than what's seen in DR1. The interferometer is located near the region of good curvature closest to the cathode.

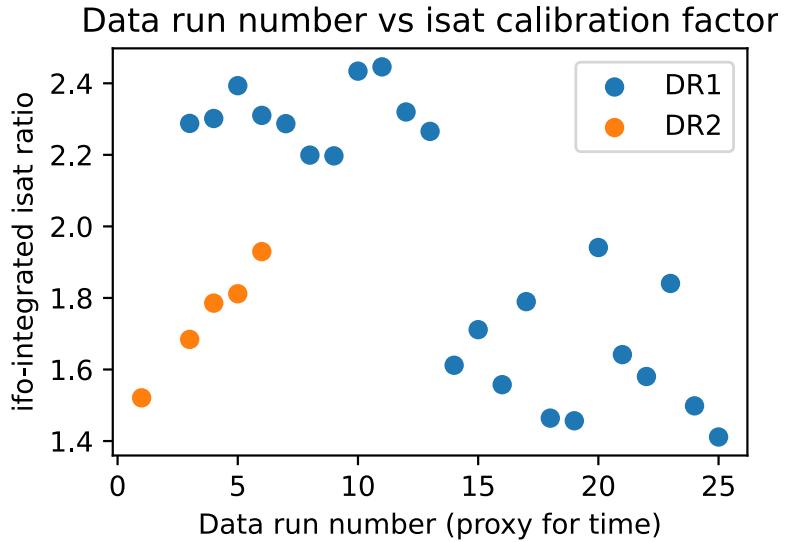


Figure 2.40: I_{sat} calibration factor over different dataruns from the same run sets. Datarun number is monotonically increasing, so in this case it's used as a proxy for time. A rather large variance is seen in this calibration factor, which suggests the I_{sat} probes having time-varying characteristics affecting the measurement.

Differences could also occur within dataruns. Calibrating the effective area of the I_{sat} probes can be done using the 56 GHz interferometer, but this calibration factor drifted over time and seen in fig. 2.40. This could be deposits being removed or added to the probe, affecting the effective area. This calls into question the reliability of absolute I_{sat} measurements, but we proceed regardless because there's no easy way to fix this issue.

2.6 Conclusions and future work

Turbulence and transport was studied in mirrors with varying lengths and ratios using the flexible magnetic geometry of the LAPD. Particle flux and fluctuation amplitudes decreased up to a factor of two when mirror ratio was increased. The primary drivers of turbulence were identified as the rotational interchange mode, caused by spontaneous rotation, and unstable drift-Alfvén waves

driven by the density gradient. The decrease in density fluctuation amplitudes can be attributed to an increase in the gradient scale length caused by the dimensionally wider plasma at the mirror midplane. Despite imposing a mirror configuration, no signs of mirror-driven instabilities were observed. The highly-collisional, GDT-like plasma produced suppressed any velocity space instabilities. The interchange growth rate was likely suppressed to an undetectable level by line-tying, in-cell electron production, and shear flow.

Future experiments in hotter regimes with the new LaB6 cathode [QGP23a] will need to be performed to evaluate the robustness of these results, particularly concerning the stabilization of curvature-induced interchange. Additionally, the source field should be matched to the mirror mid-plane field so that the plasma remains the same radius to isolate geometric effects. Simultaneous measurements using flux and/or vorticity probes and I_{sat} are needed to concretely determine if azimuthal flow shear is modified by the mirror field, and to quantify the effect of flows on rotational interchange and drift wave instability drive in general. Multiple simultaneous axial measurements of potential would enable better understanding of the axial wavenumber and identification of possible modes.

CHAPTER 3

Creating a randomized dataset for training machine learning models

3.1 Goal

3.2 The LAPD

The Large Plasma Device (LAPD)[GPL16b, QGP23b] is a basic plasma science device located at the University of California, Los Angeles. The LAPD produces up to 18m long, 1m diameter plasmas with densities up to $3 \times 10^{13} \text{ cm}^{-3}$ and temperatures up to 20 eV, though typical operation yields temperatures around 5 eV. Probes can sample virtually any point in this plasma through unique ball valves placed every 32 cm along the length of the device, enabling the collection of time series data with high spatial resolution. The discharge repetition rate is configurable between 0.25 and 1 Hz. Additionally, the LAPD has 13 independently controllable magnet power supplies to shape the geometry of the axial magnetic field. The discharge is formed by 35 cm diameter lanthanum hexaboride (LaB₆) cathode[QGP23b] and 72 cm molybdenum anode 0.5m away (-z direction) at the southern end (+z) of the device. A cartoon of the LAPD and relevant coordinate system can be seen in Fig. 4.1.

The LAPD has many experimental control parameters for various physics studies. While the device can accommodate various insertable components, this study focuses on the parameters fundamental to the operation of the main cathode. Specifically, half way between the cathode and anode are three gas puff valves: East, West, and top. The aperture, duration, and triggering of

these valves has a large impact on plasma formation. A static gas fill system also exists but it is not used in this study. The cathode-anode voltage (and consequently, discharge power) strongly influences plasma density and temperature downstream of the source. Additionally, the magnetic field configuration substantially shapes the plasma column. One crucial variable not considered in this study is the cathode temperature, as its adjustment and equilibration requires many hours, limiting dataset diversity. This combination of diagnostic coverage, high repetition rate, and extensive configurability renders the LAPD particularly suitable for machine learning studies.

Data collection was conducted in two campaigns separated by 14 months. The initial run set is designated as DR1 and the subsequent run set as DR2. These run sets are further broken down into *dataruns* which are series of discharges (“shots”) with identical operational machine parameters. A total of 67 dataruns were collected over both campaigns.

The LAPD control parameters varied in this study were the source field, mirror field, midplane field, gas puff valve voltage, gas puff duration, and discharge voltage. The magnetic field regions are labeled in fig. 4.1 and effectively control the width of the plasma relative to the cathode in their respective regions. The gas puff voltage governs gas flow rate into the chamber, though this relationship is not yet quantified, and the gas puff duration defines the piezo valve activation period. The discharge voltage is applied across the cathode and anode 10 ms after gas puff initiation. While discharge voltage correlates to discharge current (and thus power), the current depends on the machine configuration and downstream conditions and cannot be predetermined.

These machine parameters – with the exception of gas puff duration – were randomly sampled via Latin-hypercube sampling (LHS) for 44 of the dataruns. Data were then collected with these settings. Gas puff duration was reduced for the last seven runs to 20, 10, or 5 ms (see fig. 4.3 for timings relative to I_{sat} signals). The breakdown of each setting in the dataset is given in appendix 3.3, Table 3.1. The top gas puff valve was used for only the first nine dataruns of DR2 because of equipment issues.

I_{sat} measurements were acquired along $y=0$ lines (51 dataruns total) or $x-y$ grids (16 dataruns total) with spatial resolutions varying between 1.5 to 2 cm. The fixed axial locations of the probes

were 895 cm and 831 in DR1 and 1150, 1022, 863, and 639 cm for DR2 (Fig. 4.1). Six shots were recorded at each position except for the first four dataruns in DR1 with five shots each.

3.3 Data bias

Measurements over an x-y plane, constituting $\approx 64\%$ of all shots, are predominantly acquired overnight for maximal machine utilization. These longer dataruns lead to particular machine configurations being overrepresented in the dataset.

The dataset predominantly contains gas puff durations of 38 ms. Only 6 runs in the training set have gas puff durations less than 38 ms: three have 5 ms and three have 10 ms, each having mirror ratios 1, 3, and 6 but otherwise identical configurations in an attempt to see mirror-related interchange instabilities in higher-temperature, lower-collisionality regimes. The 20 ms gas puff duration case is in the test set. This sampling bias towards the 38 ms gas puff duration suggests poor model performance is to be expected in shorter gas puff regimes. The top gas puff valve was operational for only the first nine runs of DR2.

Despite the best efforts to randomize the machine configuration, imbalance in the dataset will be present because of the relatively small amount of samples for the given actuator space. The distribution of I_{sat} signals can be seen in Fig. 3.1. The I_{sat} distribution is clearly different for DR1 and DR2, with DR1 having a much flatter distribution. These distributions imply that if the model is constrained to sample from DR2 via the run set flag, then the model is expected to predict a lower I_{sat} value in general. When predicting from the model in general, performance will likely be worse for I_{sat} values $\gtrsim 11 \text{ mA/mm}^2$.

The distribution of the selected machine settings for all the dataruns is enumerated in Table 3.1. Despite the randomization of the settings of 44 dataruns, the distribution is often uneven. This unevenness is exacerbated in the test set because that is a selection of 6 out of 67 dataruns. The remaining 23 non-random dataruns also contribute to the imbalance. For example, a source field of 1 kG and discharge voltage of 112 show up disproportionately in the dataset because data were

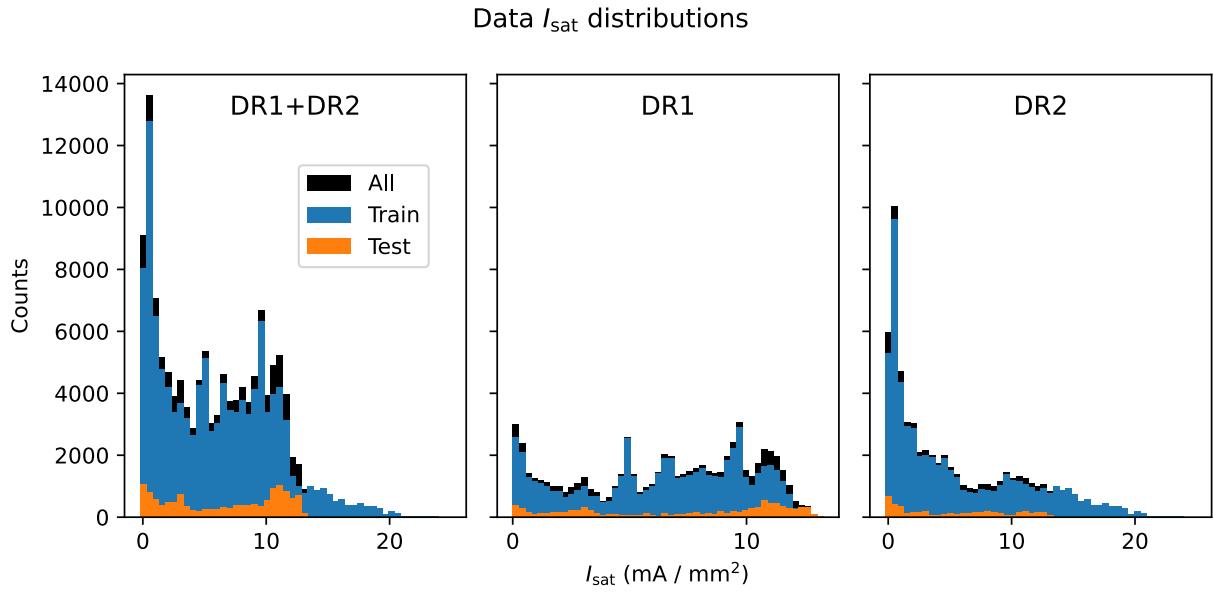


Figure 3.1: Distribution of I_{sat} signals. DR1 appears to have a more uniform distribution than DR2 does. Combining the two datasets results in many I_{sat} examples near 0 mA/mm² and a sharp decrease in number of examples above 10 mA/mm². From these histograms we expect or model to be biased towards fitting lower I_{sat} values better, and to perform badly in cases with very high I_{sat} values.

Data x-coordinate distributions, 1 cm bins

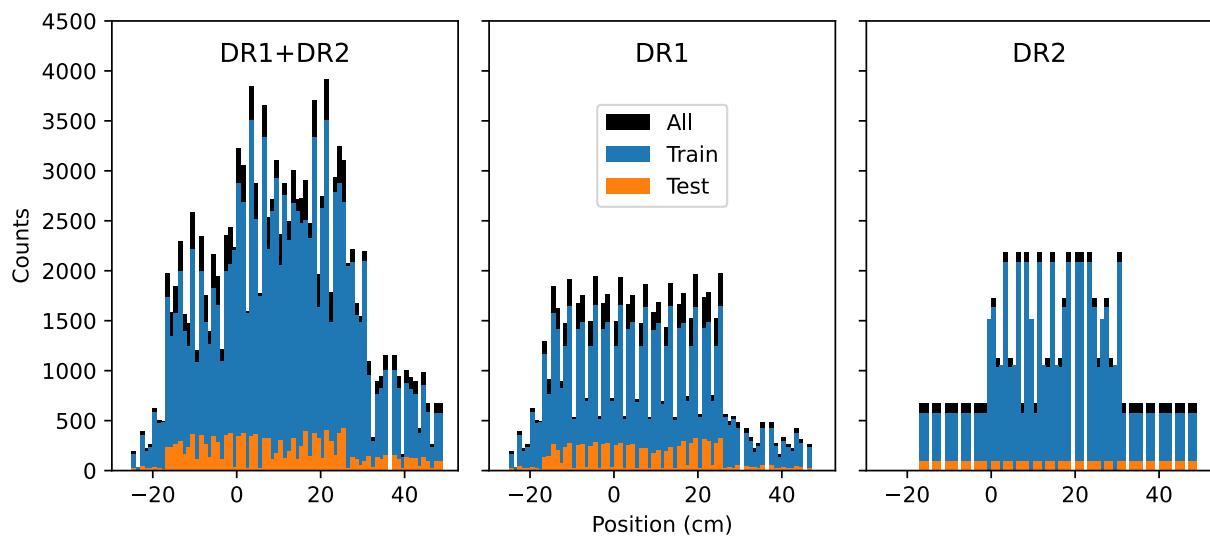


Figure 3.2: Distribution of the x-coordinate in the profiles. The increase in data points between roughly $x \approx 0$ to 30 cm is from planes instead of lines. Based on this distribution, the performance of the model is expected to be biased towards this central area.

Table 3.1: Data breakdown by class and dataset (percent)

B source (G)				B mirror (G)				B midplane (G)			
	Train	Test	All		Train	Test	All		Train	Test	All
500	4.77	0	4.29	250	4.30	8.41	4.72	250	8.25	21.01	9.55
750	3.34	12.61	4.29	500	30.49	8.41	28.23	500	43.80	8.41	40.19
1000	43.13	78.99	46.78	750	6.68	16.81	7.72	750	6.62	52.19	11.27
1250	12.59	0	11.30	1000	28.85	57.97	31.82	1000	26.36	5.78	24.26
1500	19.23	0	17.27	1250	3.34	4.20	3.43	1250	9.24	0	8.30
1750	1.91	0	1.71	1500	26.34	4.20	24.08	1500	5.73	12.61	6.43
2000	15.03	8.41	14.35								

Gas puff voltage (V)				Discharge voltage (V)				Axial probe position (cm)				
	70	12.11	16.81	70	12.22	8.41	11.83	639	12.48	8.41	12.06	
	75	6.68	0	6.00	80	5.25	0	4.72	828	17.07	36.28	19.03
	80	11.46	8.41	11.15	90	2.86	8.41	3.43	859	12.48	8.41	12.06
	82	41.49	57.97	43.17	100	3.34	8.41	3.86	895	33.01	30.10	32.71
	85	14.13	0	12.69	110	8.77	0	7.87	1017	12.48	8.41	12.06
	90	14.13	16.81	14.40	112	20.62	0	18.52	1145	12.48	8.41	12.06
					120	3.82	8.41	4.29				
					130	0.95	0	0.86				
					140	2.86	8.41	3.43				
					150	39.30	57.97	41.20				

Gas puff duration (ms)				Vertical probe position (cm)			
	38	94.27	91.59	38	≈ 0	36.26	46.08
	< 38	5.73	8.41	6.00	≠ 0	63.74	53.92

collected at those settings while other equipment was being adjusted or calibrated.

CHAPTER 4

Optimizing mirror configurations in the LAPD using machine learning

This study demonstrates the efficacy of machine learning (ML)-based trend inference using data from the Large Plasma Device (LAPD). Neural network (NN) ensembles with uncertainty quantification are trained to predict time-averaged ion saturation current (I_{sat} proportional to density and the square root of electron temperature) at any position within the dataset domain. Model-inferred trends, such as the effects of introducing mirrors or changing the discharge voltage, are consistent with current understanding. In addition, axial variation is optimized via comprehensive search over I_{sat} predictions. Experimental validation of these optimized machine parameters demonstrate qualitative agreement, with quantitative differences attributable to Langmuir probe variation and cathode conditions. This investigation demonstrates, using ML techniques, a new way of extracting insight from experiments and novel optimization of plasmas. The code and data used in this study are made freely available.

This chapter is based on the publication (currently under review) in the Physics of Plasmas titled “Machine-learned trends in mirror configurations in the Large Plasma Device” by myself, Prof. Jacob Bortnik, and Prof. Troy Carter. The goals of this work are to show what a solid, validated machine learning study looks like, and how ML can be useful in understanding operating plasma devices.

4.1 Introduction

Understanding and controlling plasma behavior in fusion devices is necessary for developing efficient fusion reactors for energy production. Because of the complex, high-dimensional parameter space, traditional experimental approaches are often time-consuming and require careful planning. This work explores how machine learning (ML) techniques can accelerate this understanding by studying the effect of machine parameters in a basic magnetized plasma device. Trend inference is this process of relationship discovery. While ML methods, particularly neural networks (NNs), have become increasingly prevalent in fusion research for control and stabilization, their application to systematic trend discovery remains largely unexplored.

Many studies have used ML for profile prediction on a variety of tokamaks, particularly for real-time prediction and control. For example, NNs were used to predict electron density, temperature, and other quantities in DIII-D [ACK21], and reservoir NNs have demonstrated the ability to quickly adapt to new scenarios or devices [JAC22]. Temporal evolution of parameters has been successfully modeled using recurrent neural networks (RNNs)[CCA24] for multiple devices, including the EAST[WYP22] and KSTAR tokamaks[SNK21, SNK22]. These predictions enabled training of a reinforcement learning-based controller[SNK21, SNK22]. In addition, a decision tree-based controller was trained to maximize β_N while avoiding tearing instabilities[FEE20] on DIII-D. Electron temperature profiles have also been predicted using dense NNs on the J-TEXT tokamak [DLD21].

A parallel focus has been on instability prediction and mitigation in tokamaks, particularly of disruptions. Notable achievements in disruption prediction include RNN-based disruption prediction [KST19] and random forest approaches[RME19], with a comprehensive review available by Vega et al [VMD22]. Recent work has extended to active control, such as the mitigation of tearing instabilities in DIII-D using reinforcement learning [SKJ24].

While ML has proven effective for prediction and control tasks, inferring trends using data-driven methods has been relatively uncommon. Notable exceptions include finding scaling laws

on the JET tokamak[MPL20] via classical ML techniques and the development of the Maris density limit[MRP24] which outperforms other common scalings (including the Greenwald density limit) in predictive capability.

The use of machine learning and Bayesian inference in fusion research has been recently reviewed by Pavone et al.[PMK23]

Outside of magnetized plasmas, the laser plasma community has embraced ML techniques for various applications, enumerated in a review by Dopp et al.[DEH23]. Data-driven plasma science in general has been reviewed by Anirudh et al.[AAA23] Notably, a similar quasi-random method (Sobol sequences) was used to collect a spectroscopy dataset on a plasma processing device over diverse machine settings [DFH23]. This process is similar to what is performed in our work here, but a generative variational autoencoder was instead trained to be used as an empirical surrogate model.

This work advances data-driven methods in plasma physics by taking these methods one step further: instead of learning a model for particular task (e.g., disruption prediction or profile prediction), we infer learned trends directly from the model itself.

The goal of this study is to develop a data-driven model that can provide insight into the effect of machine parameters on plasmas produced in Large Plasma Device (LAPD) in lieu of a theoretical model. In contrast with tokamaks and other fusion devices, the LAPD is particularly well-suited for ML data collection because of its flexibility and high repetition rate. We demonstrate the capability to infer trends in a particular diagnostic signal, the time-averaged ion saturation current (I_{sat}), for any mirror (or anti-mirror) field geometry in a variety of machine configurations. Langmuir probes are commonly used to measure density, temperature, and potential in virtually all plasma devices in low-temperature (less than 10s of an eV) regimes. The I_{sat} signal in particular is almost always used in the LAPD for calculating local plasma density.

This study performs two firsts in magnetized plasma research: using NNs to directly infer trends and collecting data efficiently with partially-randomized machine parameters. We also

demonstrate optimizing LAPD plasmas given any cost function by minimizing axial variation in I_{sat} . This global optimization is only possible using ML techniques. This work demonstrates the usefulness of a pure ML approach to modeling device operation and shows how this model can be exploited. We encourage existing ML projects and experiments to consider this approach if possible. Acquiring sufficiently diverse datasets may require assuming some risk because diverse data, such as discharges from randomly sampled machine settings, may not be amenable to conventional analysis techniques.

All the processed data used for training the models in this study are freely available[Tra25] (see section 4.11). Other devices have also made data publicly available. Namely, data for H-mode confinement scaling has been available since 2008[RWB08], and more recently some MAST[JKC24] and all LHD[lhd] data are now publicly available.

This paper is organized as follows: Section 4.2 discusses the LAPD and the data acquisition methodology. Sections 4.3 and 4.4 detail development of the model and uncertainty quantification. Section 4.5 presents model validation, followed by trends in discharge voltage and gas puff duration in Section 4.7. Section 4.8 demonstrates optimization of the axial I_{sat} profile, with the discussion and conclusion in sections 4.9 and 4.10.

4.2 Data collection and processing

4.2.1 The Large Plasma Device (LAPD)

4.2.2 Data collection and processing of I_{sat} signals

The ion saturation current, denoted as I_{sat} , is obtained by applying a sufficiently negatively bias to a Langmuir probe to ensure the exclusive collection of ions. This collected current is proportional to $S n_e \sqrt{T_e}$, where n_e and T_e are the electron density and temperature, and S is the effective probe collection area. To account for differences in probe tip geometry, the I_{sat} values are normalized to area. Under typical conditions, an I_{sat} value of 1 mA/mm² corresponds to $n_e \approx 1.2 \times 10^{12} \text{ cm}^{-3}$

Large Plasma Device

Top view

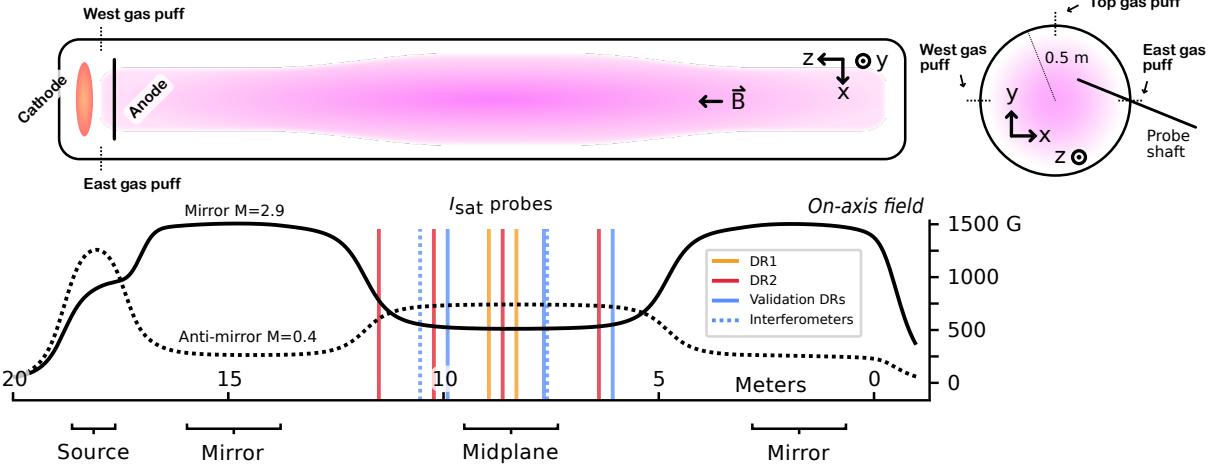


Figure 4.1: A cartoon of the Large Plasma Device, the coordinate system used, examples of a mirror and anti-mirror magnetic field configuration, and probe locations used in this study. The source, mirror, and midplane regions are labeled; the three fields were programmed independently.

for a T_e from 4 to 1 eV.

I_{sat} measurements were averaged over 10 to 20 ms to exclude plasma ramp-up and fluctuations. Example I_{sat} probe data can be seen in fig. 4.3 along side gas puff timings. For the probe tip that was on the same shaft as the swept probe (in DR2), the signal was instead averaged over when the bias voltage on the swept tip was held constant at the lowest value. A 40 μs (250 sample) buffer was used after the sweep was turned off to minimize the impact of transient conditions. A comparison of the full trace and the trace with the swept portion excluded can be seen in fig. 4.2. Notably, the measured I_{sat} value does not attain a steady state before the discharge shuts off.

Profile evolution is not studied to minimize computational requirements. I_{sat} characteristics vary significantly between axial (z) position machine parameters. For I_{sat} measurements on the same probe as a Langmuir sweep (DR2 port 26, $z=863$ cm), the averaging process excludes the sweep period with an additional 40 μs buffer. I_{sat} measurements in DR1 that saturated either the isolation amplifier or digitizer are excluded from the dataset. Only 484 shots were removed out of

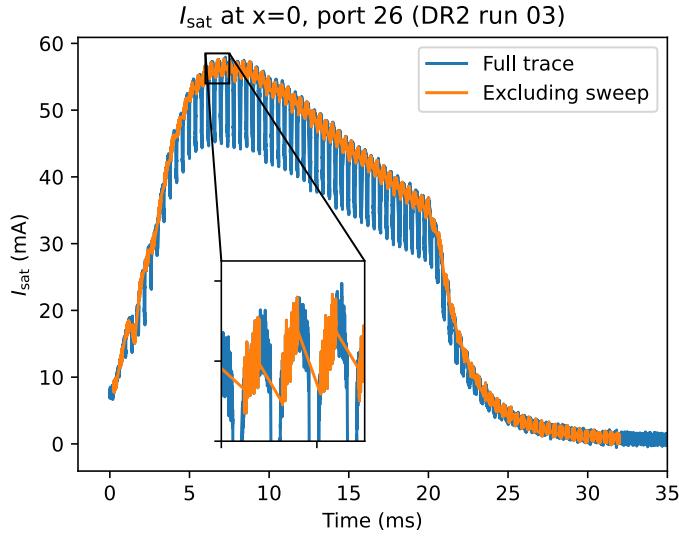


Figure 4.2: I_{sat} traces from the swept probe (port 26) from DR2 datarun 03, shot 1 of 6. The orange curve is excluding times when a sweep is active on an opposing tip.

$\approx 132,000$, so the impact on the aggregate dataset is minimal.

While I_{sat} exhibits a small degree of shot-to-shot variation, the present model only learns the expected value, leaving distributional modeling to future generative approaches. An example of these I_{sat} profiles and the six-shot variance can be seen in fig. 4.4.

4.2.3 Dataset distribution and bias

Neural network inputs comprise 12 variables: source field, mirror field, midplane field, gas puff voltage, discharge voltage, gas puff duration, probe coordinates (x, y, z), probe rotation, run set identifier, and top gas puff flag. These variables can be interpreted as six control parameters, four probe coordinates, and two flags. These inputs are mean-centered and normalized to the peak-to-peak value with no outliers in the dataset. The baseline models trained in section 4.3.3 did not contain the run set identifier or top gas puff flag.

The I_{sat} distribution is dissimilar between DR1 and DR2: DR1 appears to have a more uniform

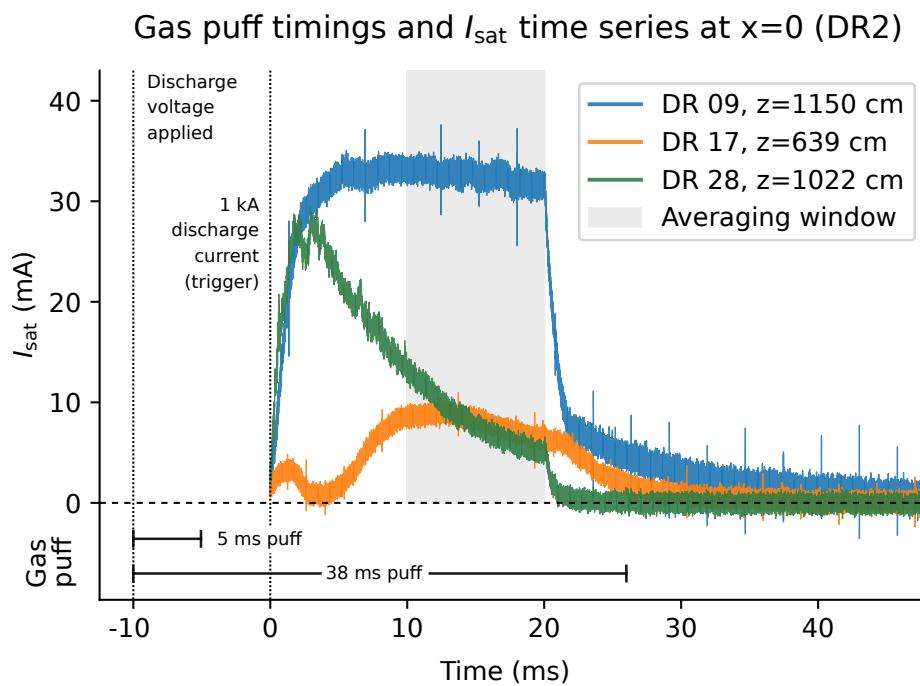


Figure 4.3: Gas puff timings and example I_{sat} time series at three different z-axis locations from three different dataruns. Note that some discharges do not achieve steady state in I_{sat} .

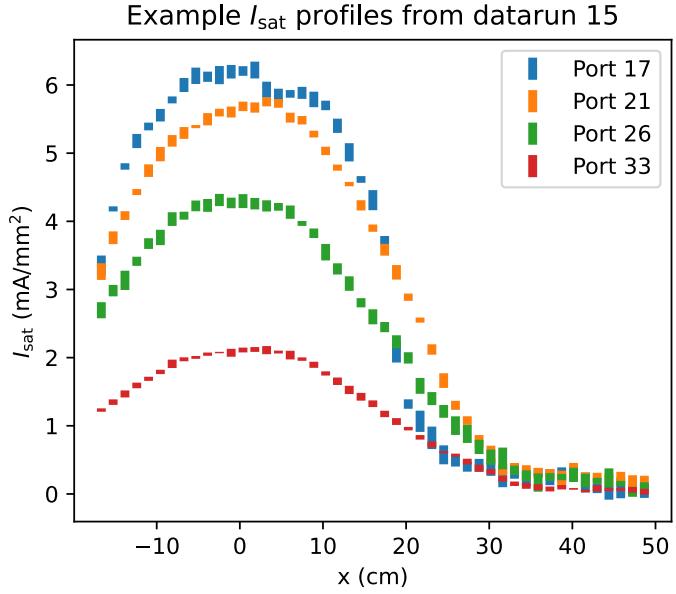


Figure 4.4: Examples of I_{sat} profiles from DR2 run 15. The bars represent the minimum and maximum of the six I_{sat} measurements taken at that position.

distribution. Combining the two datasets results in many I_{sat} examples less than 2 mA/mm^2 and a sharp decrease in number of examples above 10 mA/mm^2 . Thus, we expect the model to perform better for smaller I_{sat} values than larger ones. Data bias is further discussed in appendix 3.3.

4.3 Model development, training

4.3.1 Training details

For initial experiments in training the model, a mean-squared error (MSE) loss is used:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{m} \sum_{i=1}^m (f(x_i) - y_i)^2 \quad (4.1)$$

where x_i represents the input vector for the i th example, y_i the target measurement, m the batch size, and f the NN. During training, overfitting was assessed via the validation set MSE with a traditional 80-20 train-validation random split. Unless stated otherwise, a dense neural network, 4

hidden layers deep and 256 units wide (201,218 parameters for β -NLL loss, 200,962 parameters for MSE loss), was trained with AdamW using a learning rate of 3×10^{-4} . Leaky ReLU activations (the nonlinearities in the NN) and adaptive gradient clipping[SWP20] (cutting gradients norms above the 90th percentile of recent norms) were used to mitigate vanishing gradients and mitigate exploding gradients, respectively. The models were evaluated after training concluded at 500 epochs.

4.3.2 Validating the training pipeline

ML training processes are relatively simple but bugs, particularly in the data pipeline, can be insidious and can affect final model performance even though training looks fine. Here we validate the data pipeline (which should be performed in every ML study) to verify that the model is training and expected and that there is no accidental data leakage between the train and test sets. Andrej Karpathy’s advice for training neural networks [Kar19] was used as a template for verifying the training procedure used in this project. The data fed into the model immediately before the forward pass (and subsequent backpropagation) was stored and verified: the data are correctly randomly shuffled in each batch. Each epoch contains the same random shot order. To validate the data pipeline, a simple dense model (4 layers, 512 wide with one output; 794113 parameters, tanh activations) was trained. The model is also overfit on a single batch (128 examples) of training data to make sure that training progresses as expected. A deep double descent is observed as expected [NKB19, SKR23]. Training on a batch of 8 examples reaches ≈ 0 training loss after 50 steps. Plots of the train and validation losses can be seen in Fig. 4.5.

Multiple models were trained with varying depths and widths to verify that training loss decreases with increased model capacity. Doubling the layer width from 512 to 1024 moderately decreases the training loss; doubling the depth of the network from 4 to 8 layers has a larger impact. Increasing the width further to 2048 and depth to 12 layers has a dramatic impact on training loss, so this model and dataset are behaving nominally. The model pipeline is training and performing as expected, so we proceed.

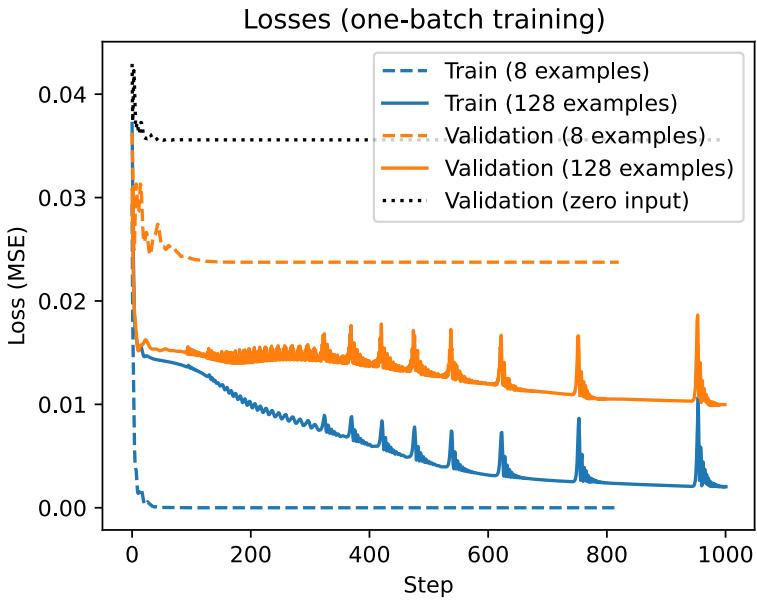


Figure 4.5: Training and validation losses when overfitting the model. A deep double descent in the validation losses is observed when fitting a single batch of 128 examples. The 8-example batch hits near-zero loss after 50 steps. This process verifies our training process is functioning as expected. The spikes are from exploding gradients which can be mitigated by clipping the gradients. A model trained on blank data is also shown as the black dotted line.

4.3.3 Baselines for mean-squared error

A model was first trained with zeroed-out inputs as a baseline and to validate the data pipeline. This model effectively has only a single, learnable bias parameter at the input. This process yields a validation loss (simply MSE in this case) of 0.036.

A linear model obviously cannot fit the dataset (see the nonlinear shape in fig. 4.4). However, a simple (and mostly linear) model can provide a performance baseline to help spot bugs when training more complex models. Since the x- and y- profiles have an approximate tanh shape, a feature is added at the linear model input stage for the x and y coordinates: $x_{\text{tanh}} = c \cdot \tanh(|x + s| \cdot a + b)$ where s, a, b, c are trainable parameters (independent for each coordinate; c is superfluous). This function was chosen to give the linear model the capability of expressing tanh-like curves. The performance of the linear model on DR2 data, with and without the tanh features, can be seen in fig. 4.6. This baseline linear-like model reaches a training and validation loss of around 0.011, with the RMSE = $\sqrt{\text{loss}} \sim 0.1$. The linear-only model is marginally worse with losses at around 0.014.

This feature engineering-like approach can continue. For example, the width of the profile is largely controlled by magnetic field configuration of the device, particularly by $\sqrt{B_{\text{midplane}}/B_{\text{source}}}$. This behavior can be added to this model, either as a new feature or as a custom relationship in the model. Note that, as seen in fig. 4.6, the width of the profile also depends on the axial coordinate. Combined with other coordinates and actuators, like discharge voltage and gas pressure, the number of possible features or function space grows combinatorially, making this custom fitting process difficult and tedious to design and test by hand. The obvious solution would be to use symbolic regression or fitting to a function library which may be ideal methods if simple profile prediction were the final goal. However, we are ultimately interested inferring trends in a much more complex input space where neural networks are more flexible and accurate. If NNs do face generalization issues, symbolic regression or a SINDy-like approach can be used instead, albeit with limited applications. Symbolic methods are appealing because the fits are simple. However, even though a simple equation may fit the data well, it does not necessarily provide insight or relate to

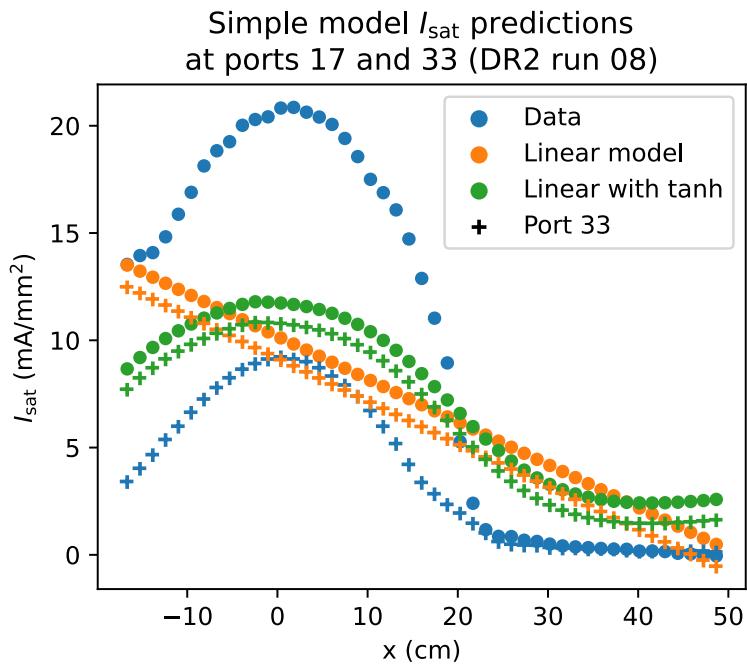


Figure 4.6: I_{sat} profiles and predictions for ports 17 and 33 based on inputs from DR2 run 08 using a liner and linear-with-tanh models. DR2 run 08 is in the training set. The “data” points are averaged over six shots. Run 08 was chosen for its representative performance; ports 17 and 33 were chosen to demonstrate the maximal axial variation (across 511 cm). These models fail to describe the data accurately.

Table 4.1: Summary of test set losses for different training data and ensembles

Model	$MSE \times 10^{-3}$
Zeroed-input	36 (validation)
Linear model	14 (validation)
Linear with tanh features	11 (validation)
9 dataruns	7.0
19 dataruns	6.9
29 dataruns	4.2
39 dataruns	4.1
49 dataruns	3.4
DR1 only	6.4
DR2 only	5.4
Full set, large model	2.8
Full set average	3.6 ± 0.56
Full set ensemble	2.9 ± 1.1
“Run set” flag average	2.1 ± 0.15
“Run set” flag ensemble	1.9 ± 0.64
“Top gas puff” flag	1.8

the underlying physics; using a freeform fitting function like a neural network is more appropriate in this use case.

A summary of these baselines is seen at the top of Table 4.1.

4.3.4 Effects of training set and model sizes

To study the effects of reduced diversity, the number of unique dataruns in the training set was systematically reduced while evaluating on a fixed test set. The test set loss monotonically increased with this decrease in datarun count. Part of this decrease may be caused by a simple reduction in

training set size. In addition, models were individually trained and evaluated on DR1 only or DR2 only. When evaluated on the left-out run set, the test set losses were high, near or above the zero-input baseline of 3.6×10^{-2} . This result suggests that both run sets contain significant information missing in the other, and training on both provides beneficial information on the structure of the I_{sat} measurement despite different probe calibrations and cathode state.

A larger model, consisting of a 12-deep 2048-wide dense network, was trained on the full training dataset, evaluated at 30 epochs. This larger model yielded a test MSE of 2.8×10^{-3} , indicating that these NNs are behaving as expected. Longer training or larger models may yield better test set results, but will likely not come close to the training and validation losses which are on the order of 10^{-5} . Combined models with differing initializations (an ensemble), were trained to measure the MSE variance over model parameters which was about 16%. When the I_{sat} predictions were averaged, the test set MSE was $2.9 \pm 1.1 \times 10^{-3}$, achieving the best performance for that model size. These test set losses are also seen in Table 4.1.

4.3.5 Improving performance with machine state flags

Data from DR1 and DR2 were collected 14 months apart leading to differing machine states. In DR1, only one turbo pump was operating leading to much higher neutral pressures than in the DR2 run set. A new parameter (mean-centered and scaled) was added to the inputs to distinguish between these two run sets. All the predictions in this study use the DR2 run set flag (a value of 1.0) because turning off the turbopumps is not a commonly desired mode of operating the LAPD. The inclusion of this parameter also provides the model the ability to distinguish between the probe calibration differences between DR1 and DR2. An ensemble prediction with this run set flag brings the test set MSE down to 1.9×10^{-3} .

A flag indicating when the top gas puff valve was enabled in DR2 was also added to all training data, allowing the model to further distinguish between different fueling cases. The addition of this flag incrementally improved test set MSE to 1.8×10^{-3} . The effect on MSE on the inclusion

of these new parameters is compared to the performance of other models in Table 4.1.

4.3.6 Learning rate scheduling

Modifying the learning rate over time (scheduling) is known to improve model learning. The following schedules were compared: constant learning rate ($\gamma = 3 \times 10^{-4}$), $\gamma \propto \text{epoch}^{-1}$, $\gamma \propto \exp(-\text{epoch})$, and $\gamma \propto \text{epoch}^{-1/2}$. The epoch is the training step divided by the number of batches in one epoch, so “epoch” in this case takes on a floating-point value. $\gamma \propto \text{epoch}^{-1}$ appears to give the best test set loss by a test MSE difference of 1×10^{-4} , and any schedule beats a constant learning rate by a difference of $2 - 4 \times 10^{-4}$.

4.4 Uncertainty quantification

4.4.1 β -NLL loss

Instead of predicting a single point, the model can predict a mean μ and variance σ^2 using the negative-log likelihood (NLL) loss [NW94, LPB17] by assuming a Gaussian likelihood. An adaptive scaling factor $\text{StopGrad}(\sigma_i^{2\beta})$ is introduced that can be interpreted as an interpolation between an MSE loss and Gaussian NLL loss, yielding the β -NLL loss:

$$\mathcal{L}_{\beta-\text{NLL}} = \frac{1}{2} \left(\log \sigma_i^2(\mathbf{x}_n) + \frac{(\mu_i(\mathbf{x}_n) - y_n)^2}{\sigma_i^2(\mathbf{x}_n)} \right) \text{StopGrad}(\sigma_i^{2\beta}) \quad (4.2)$$

for example n and model i , with an implicit expectation over training examples. $\beta = 0$ yields the original Gaussian NLL loss function and $\beta = 1$ yields the MSE loss function. This factor improves MSE performance by scaling via an effective learning rate for each example (which necessitates the `StopGrad` operation) [STA22], and improves both aleatoric and epistemic uncertainty quantification [VS22]. $\beta = 0.5$ was used by default in this study. This β -NLL loss function also improved training stability.

This NLL-like loss assumes the prediction – the likelihood of y given input \mathbf{x} : $p(y|\mathbf{x})$ – follows a Gaussian distribution. Treating each prediction as an independent random variable (considering each model in the ensemble is sampled from some weight distribution $\theta \sim p(\theta|\mathbf{x}, y)$) and finding the mean of the random variables yields a normal distribution with mean $\mu_*(\mathbf{x}) = \langle \mu_i(\mathbf{x}) \rangle$ and variance $\sigma_*^2 = \langle \sigma_i^2(\mathbf{x}) + \mu_i^2(\mathbf{x}) \rangle - \mu_*^2(\mathbf{x})$ where $\langle \rangle$ indicates an average over the ensemble.

The loss function for one of the NNs in an ensemble is seen in fig. 4.7. The MSE decreases monotonically for the training and validation set, but does not for the test set. The loss function can no longer be interpreted as a log-likelihood because of the effective per-example learning rate set by the β term in the loss (eq. 4.2). Note that early stopping (at around 8 epochs) would improve the test set loss, but the MSE would still be several factors higher than after 500 epochs. Early stopping was not explored in this study.

The ensemble predictive uncertainty can be broken down into the aleatoric and epistemic components [VS22]: the aleatoric uncertainty is $\langle \sigma_i^2(\mathbf{x}) \rangle$ and the epistemic uncertainty is $\langle \mu_i^2(\mathbf{x}) \rangle - \mu_*^2(\mathbf{x}) = \text{Var}[\mu_i(\mathbf{x})]$. The intuition behind these uncertainties is that the random fluctuations in the recorded data are captured in the variance of a single network, σ_i^2 . If the choice of model parameters were significant, we would expect the predicted mean for a single model, μ_i , to fluctuate as captured by $\text{Var}[\mu_i(\mathbf{x})]$.

4.4.2 Cross-validation MSE

For cross-validation, multiple train-test set pairs were created. Test set 0 comprises deliberately chosen dataruns to encompass a diverse set of machine settings and probe movements. The other six datasets were compiled with randomly chosen dataruns (without replacement) while keeping the number of dataruns from DR1 and DR2 equal. Seven model ensembles (5 NNs per ensemble – 35 NNs total) were trained to evaluate the effect of test set choice on model MSE. The test set MSE performance can be seen in fig. 4.8, and the training MSE performance in fig. 4.9. The median ensemble test set MSE for these seven sets was 2.13×10^{-3} with a mean of 3.6×10^{-3} .

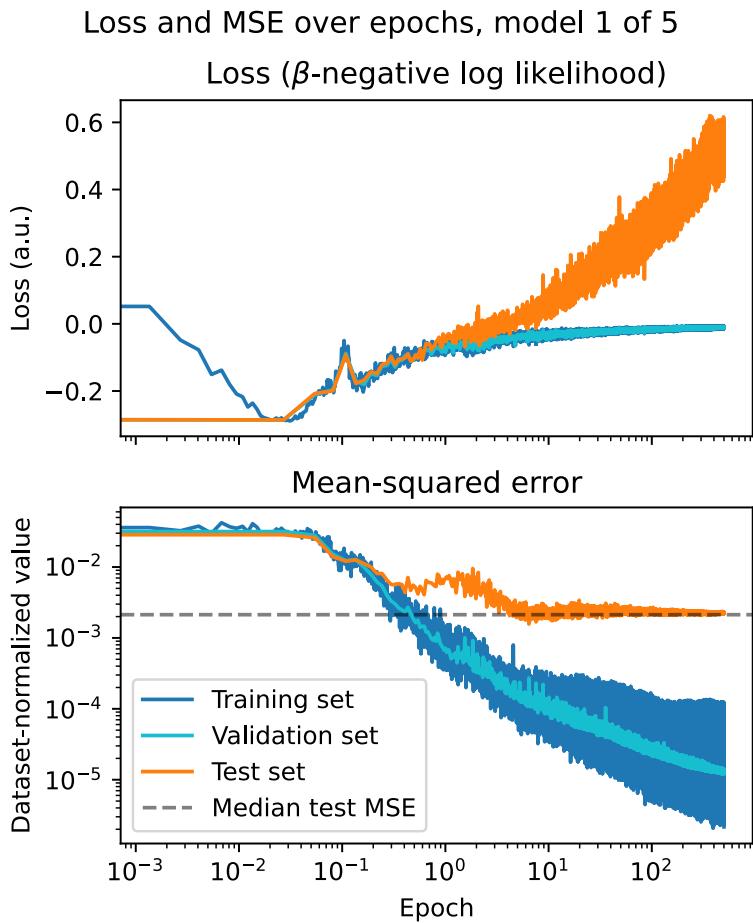


Figure 4.7: The loss and MSE for the training, validation, and test sets over the entire training duration of 500 epochs. The inclusion of the β term in the loss function – interpreted as a per-example learning rate – makes the loss function no longer interpretable in simple terms. The mean-squared error benefits from longer training for all sets.

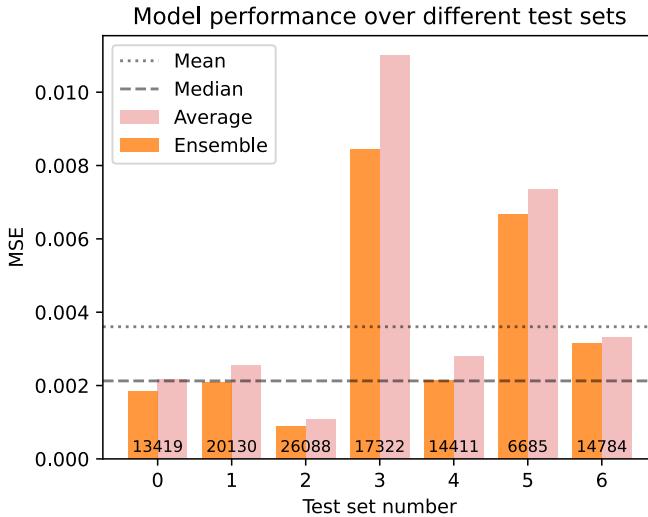


Figure 4.8: Model performance as measured by MSE over test sets with different dataruns. Test set 0 is the hand-picked dataset, and the rest were randomly compiled without replacement (though separate for DR1 and DR2). The number at the bottom of the bar chart is the number of shots in the testing set. The median test set performance is very close to the hand picked (set 0) performance. Ensembles always out-perform the average single-model prediction.

The handpicked dataset had an ensemble test set MSE of 1.85×10^{-3} , indicating that the choice of dataruns was adequately representative. This median MSE will be used to estimate model prediction error in addition to uncertainty quantification. This cross-validation also provides an error estimate if the models were to be trained on *all* dataruns. Ensembles always out-performed the average error of single-model predictions.

All validation set MSEs fall between 1 and 6×10^{-5} , with the average training MSE falling within that range as well. These MSEs indicate that the model is able to fit the training data to a high degree of accuracy regardless of which dataruns are held out. The loss and MSE curves over training epochs can be seen in the appendix in fig. 4.7.

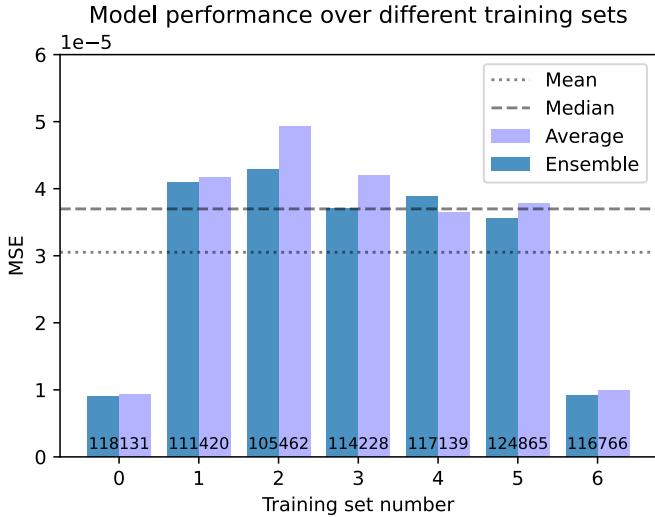


Figure 4.9: See caption for fig. 4.8. Note that the training loss is dramatically less than the testing loss, but otherwise there is no discernible trend.

4.4.3 Model calibration via weight decay

The predicted uncertainty may not provide an accurate range of I_{sat} values when compared to the measured value. Calibrating the model means changing the predicted uncertainty range so that the measured values fall within that range according to some distribution, such as a Gaussian in this case. One of the ways assessing this calibration is by the z-score of predictions, where $z_n = (x_n - \mu_n) / \sigma_n(x_n)$ for example x_n , predicted mean μ_n , and standard deviation σ_n . Perfect model calibration would lead to identical z-score distribution $\mathcal{N}(\mu = 0, \sigma = 1)$ for the training and test sets. When evaluated on the training set, the distribution should be a Gaussian with a standard deviation of 1. The z-score distributions for the train and test sets with a model weight decay of 0 can be seen in fig. 4.10.

Increased weight decay can lead to better model calibration [GPS17]. Weight decay penalizes large parameter values by adding the L2 norm of model weights to the loss. Model ensembles were trained with weight decay coefficients between 0 and 50 to determine the best calibrated model determined by the distribution of z-scores of the training and test sets. The results of this weight

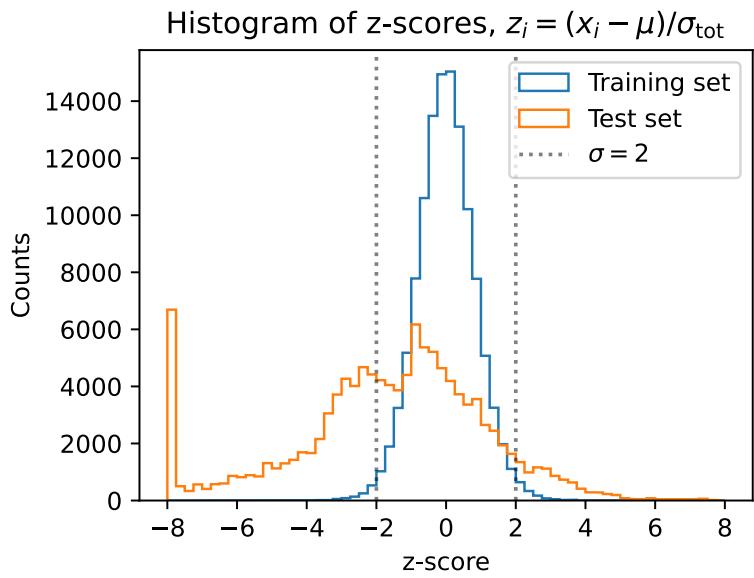


Figure 4.10: Z-scores for the training and testing set for the model with a weight decay coefficient λ of zero. The magnitude of counts for the test set is scaled up by a factor of 8.8 (the train-to-test example ratio). The histograms are clipped between z of -8 and 8 with a bin width of 0.25; the spike at the negative side of the test set histogram is from the long tail.

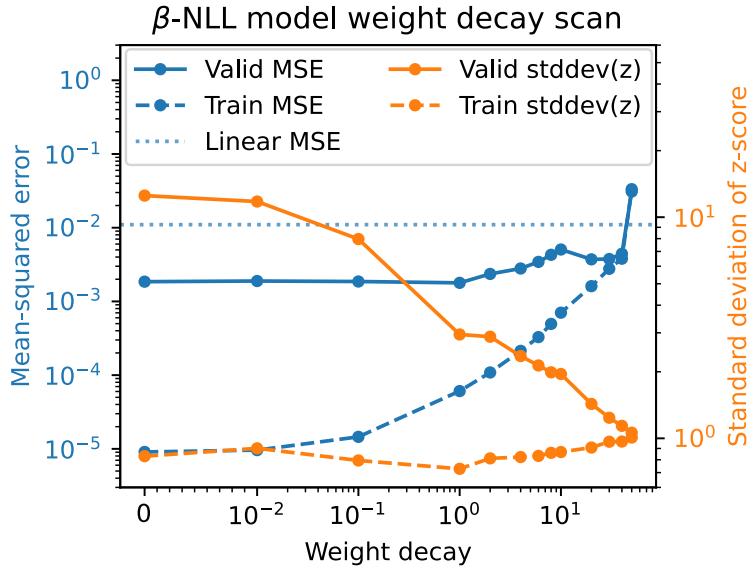


Figure 4.11: Model performance and calibration for different weight decays. Highly biased models are better calibrated, but come at great expense of mean prediction error. At the weight decay value of 50, the model has worse error than a linear model. Note the linear scale below 10^{-2} .

decay scan are seen in Fig. 4.11. Increasing the weight decay increases the test MSE and decreases its z-score standard deviation. This large standard deviation is caused by outliers. Excluding z-scores magnitudes above 10, or 4.4% of the test set, yields a standard deviation of 2.53. This long tail indicates that the distribution of predictions on the test set is not Gaussian. Nonetheless, the trend remains that increasing weight decay leads to smaller test set z-score standard deviations. However, the test set MSE increases after a weight decay of 1. This increase in test MSE implies that the model is making less accurate predictions but is better calibrated. Highly biased models are better calibrated, but come at great expense of mean prediction error. At the weight decay value of 50, the model has worse error than a linear model. Despite the attempts using weight decay, the model never becomes well-calibrated: the predicted uncertainty is always too low by a factor of 2 to 5.

Despite the better calibration, the uncertainty predicted by a model with a large weight decay

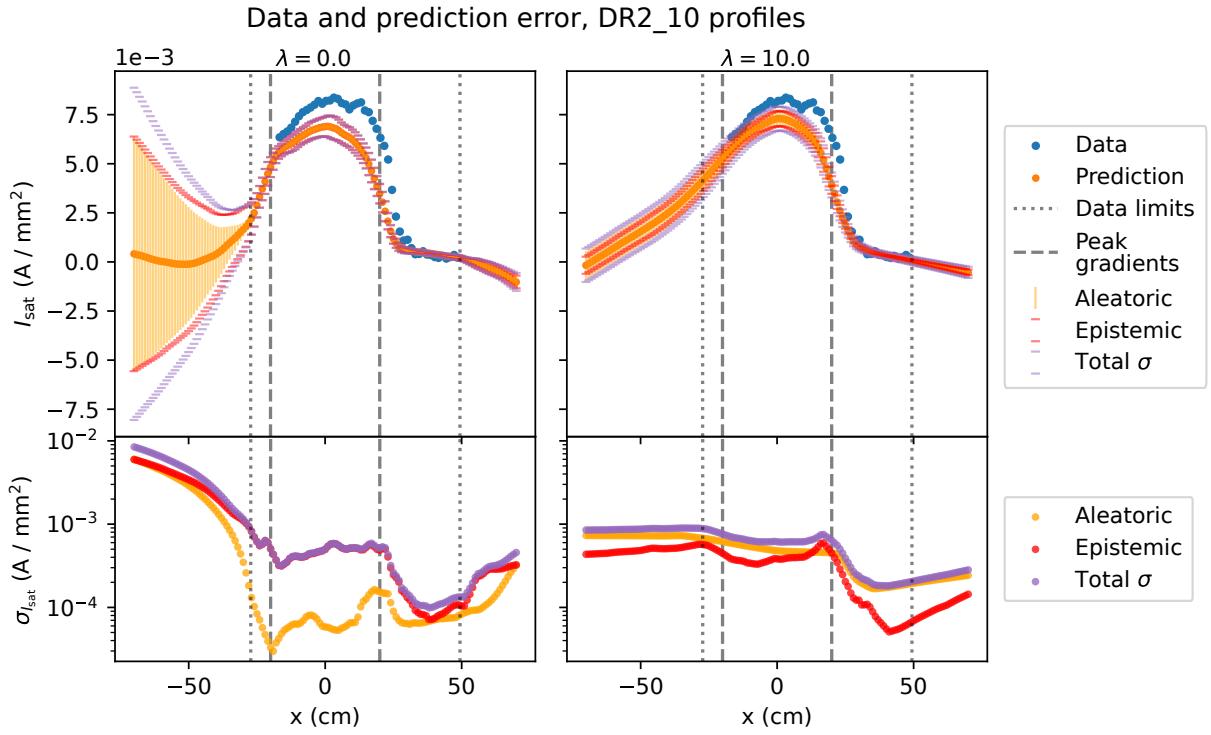


Figure 4.12: Model extrapolation performance (top plots) with uncertainty (bottom plots) for a model ensemble trained on a β -NLL loss function. DR2 run 10 was chosen as an illustrative example. The *relative* uncertainty appears to be more useful when zero weight decay ($\lambda = 0$, left) is used: the uncertainty increases when the model is predicting outside its training data along the x-axis.

is decidedly worse: the uncertainty is similar across an entire profile, and when projected beyond the training data, the total uncertainty remains largely constant as seen in Fig. 4.12. The zero weight decay model exhibits relatively increasing uncertainty beyond the bounds of the training data. Although not well-calibrated, this uncertainty can provide a hint of where the model lacks confidence relative to other predictions, even though the uncertainty is much less than it should be.

4.5 Evaluating model performance

Model performance is evaluated in three ways by comparing against intuition from geometry, an absolute measurement, and extrapolated machine conditions.

4.5.1 Checking geometrical intuition

Assuming magnetic flux conservation, we know that modifying the mirror geometry can control the effective width of the plasma. One way to check that the model is learning appropriate trends is to check with this intuition. If the magnetic field at the source is not equal to the field at the probe, the probe will see the plasma expanded (or contracted) by roughly a factor of $\sqrt{B_{\text{probe}}/B_{\text{source}}}$. The cathode is about 35 cm in diameter, so a magnetic field ratio of 3 would give produce a plasma approximately 60 cm in diameter. All the probes used in this study are in or very close to the zero-curvature midplane region of a mirror.

To check this intuition, the model is given the following inputs: $B_{\text{source}}=500$ G, $B_{\text{mirror}}=1500$ G, $B_{\text{midplane}}=500$ G, discharge voltage=110 V, gas puff voltage=70 V, gas puff duration=38 ms, run set flag=DR2 and top gas puff=off. The discharge voltage and gas puffing parameters were arbitrarily chosen. The x coordinate is scanned from 0 to 30 cm, and the z coordinate from 640 to 1140 cm. This discharge is then modified by separately changing B_{source} to 1500 G and B_{midplane} to 750 G ($M=1.5$). The x profiles at the midplane ($z=790$ cm) of the reference $M=3$ prediction, source field change, and midplane field change, all scaled to cathode radius, can be seen in Fig. 4.13. Changing the source field to 1500 G increases the I_{sat} towards the edge of the plasma, as expected. When the midplane field is increased, the I_{sat} values decrease at the edge and increase at the core ($x=0$ cm), implying a thinner plasma column and is consistent with previously measured behavior. When only the mirror field is modified (not shown), the strongest effect on I_{sat} is on or near $x=0$ cm, and the plasma column width does not appear to appreciably change.

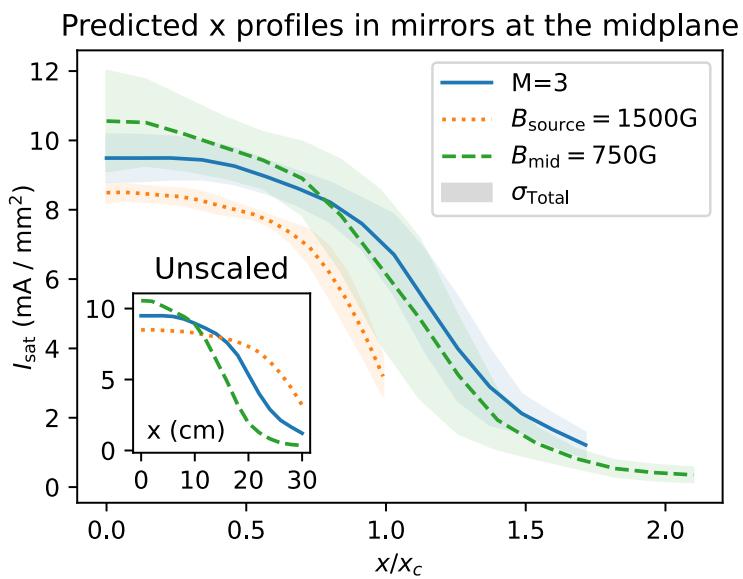


Figure 4.13: Plot of various mirror configurations scaled to the cathode radius $x_c = 17.5$ cm at the midplane ($z=790$ cm). When scaled according to the expected magnetic expansion, the profiles generally agree. The smaller the plasma diameter (and thus smaller volume), the higher the peak in I_{sat} at the core, as expected.

4.5.2 Directly comparing prediction to measurement

I_{sat} measurements were taken with the following LAPD machine settings: $B_{\text{source}}=1250 \text{ G}$, $B_{\text{mirror}}=500 \text{ G}$, $B_{\text{midplane}}=1500 \text{ G}$, discharge voltage=90 V, gas puff voltage=90 V, gas puff duration=38 ms, run set flag=DR2 and top gas puff=off. These settings were from a previous discharge optimization attempt. The probes utilized were the permanently-mounted 45° probe drives. These probes were known to have identical effective areas relative to each other from the previous experiment and from analyzing the discharge rampup.

Because of data acquisition issues, only a single useful shot was collected at a nominal -45° angle (relative to the x-axis) 10 cm past the center ($x=0 \text{ cm}$, $y=0 \text{ cm}$) of the plasma on three probes at z-positions of 990, 767, and 607 cm (ports 22, 29, and 34, respectively). The probe drives were slightly uncentered, leading to the real coordinates of the probes to be around $x = 9.75 \text{ cm}$ and $y = -8.4 \text{ cm}$. Note that the model can predict anywhere in LAPD bounded by the training data, so off-axis measurements are not an issue. The resulting predictions using these coordinates and machine conditions can be seen in Fig. 4.14. The model reproduces the axial trend well, but slightly underestimates I_{sat} on an absolute basis. However, given the lack of absolute I_{sat} calibration and variable machine state, the agreement of the absolute I_{sat} values may be coincidental. Nevertheless, the trend exhibited by this validation study match the predicted trend and increase our trust in model predictions.

An additional validation datarun was performed. For this run, the discharge voltage was increased to 160 V, and the source field changed to 822 G. The training data contains discharge voltages up to 150 V, so this case tests the extrapolation capabilities of the model. The comparison of model predictions and the measured data can be seen in Fig. 4.14. As stated earlier, the absolute uncertainty provided of the model is not calibrated. However, note that the level of uncertainty provided by the model, as well as the large spread in model predictions, are much greater than seen in the interpolation regime (Fig. 4.14) and eclipses the cross-validated test set root mean squared error (RMSE). We conclude that this model has good interpolation capabilities, but extrapolation

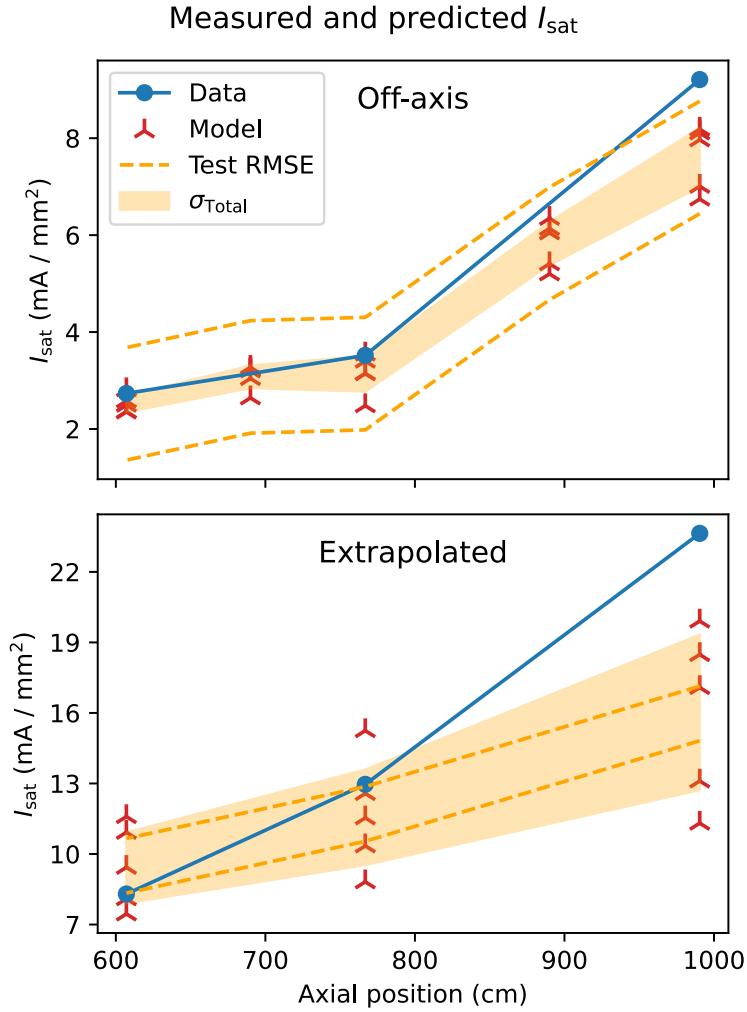


Figure 4.14: Top: data collected at off-axis positions around $x = 9.75$ cm and $y = -8.4$ cm are compared with predictions from the machine learning model at the same points in addition to two interpolating predictions. The model predicts the trend well, but underestimates I_{sat} in general. The shaded orange region is the total model uncertainty ($\sigma = \sqrt{\text{Var}}$). Bottom: Measured vs predicted I_{sat} values for an odd machine configuration with $B_{\text{source}}=822$ G and discharge voltage=160 V. The training data only covers discharge voltages up to 150 V. The machine was also in an odd discharge state, so it's no surprise that the predicted uncertainty bounds are very large (even greater than the test set RMSE value) and that accuracy suffers.

– as with any model – remains difficult.

4.5.3 Comparison with Thomson scattering

The z-axis interpolation for dataruns in the training and test sets can be evaluated using the Thomson scattering (TS) diagnostic. The TS measurement is taken 8 ms into the discharge for DR1 12 ms into the discharge for DR2, but in this study the measured and predicted I_{sat} are instead averaged over 10 to 20 ms. The Thomson scattering measurement is compared in with DR1 and DR2 in figs. 4.16a and 4.16b, respectively. The linear slope fits do not take model error into account. In DR1, I_{sat} predictions disagree with the I_{sat} derived from TS. Measurements from probes, when nearby the TS beam, can also have very different values from the TS-derived measurement. The TS density measurement may suffer from misalignment, and has not been calibrated since January 2022, roughly a year (DR1) or two (DR2) before these data were taken. The density measurement is photon counting and requires absolute calibration. This disagreement likely comes from this error in density because fitting the temperature is robust to absolute calibration errors. In addition, the I_{sat} is time-varying; the average may differ substantially from single points in time earlier in the discharge.

All these issues considered, the model predictions has rough agreement with TS on average in DR1, which is encouraging because the TS beam at port 32 (671 cm) is substantially further from the closest probe at port 27 (831). We should expect rough agreement or a slight underestimate on average based on the skewed test-set z-score seen in fig. 4.10). DR2 has a probe past the TS beam at port 33 (639 cm), but the I_{sat} measurement rarely agrees with Thomson. Because of this density error and measurement time discrepancy, we conclude that the TS diagnostic may not be a good way to verify the predictions of the model. Note that, when calibrated, TS agrees with I_{sat} measurements quite closely as seen in the LAPD Thomson scattering paper [GKG22].

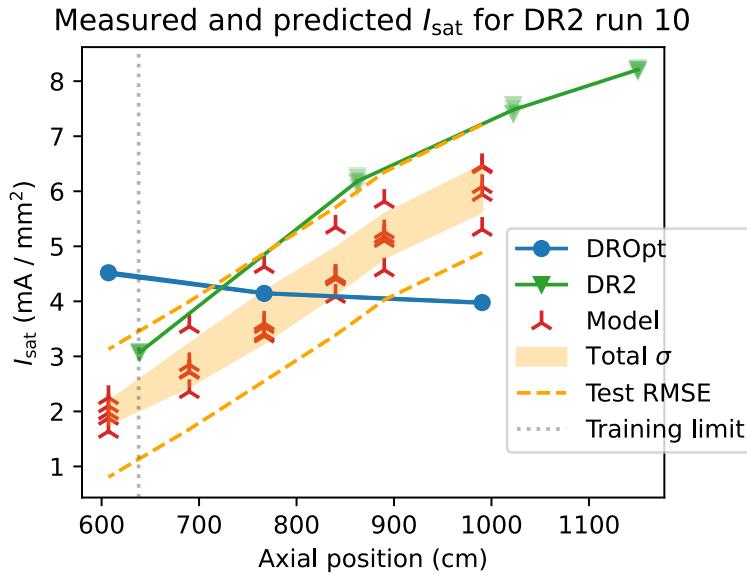
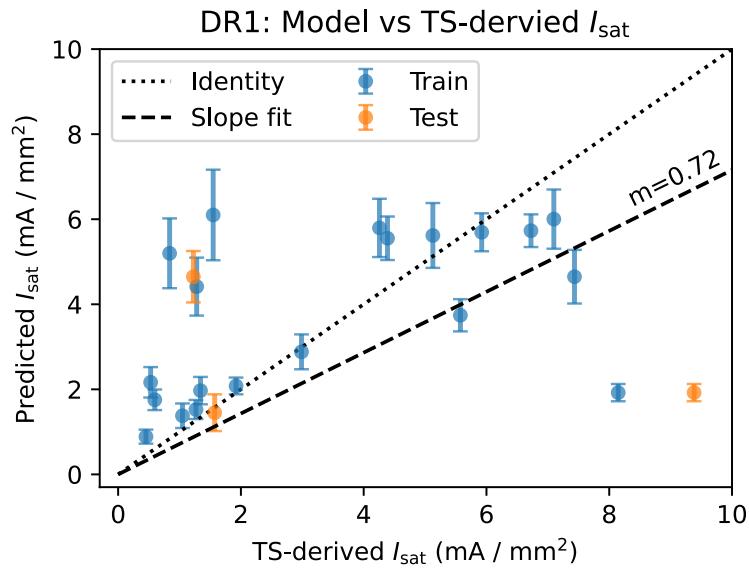


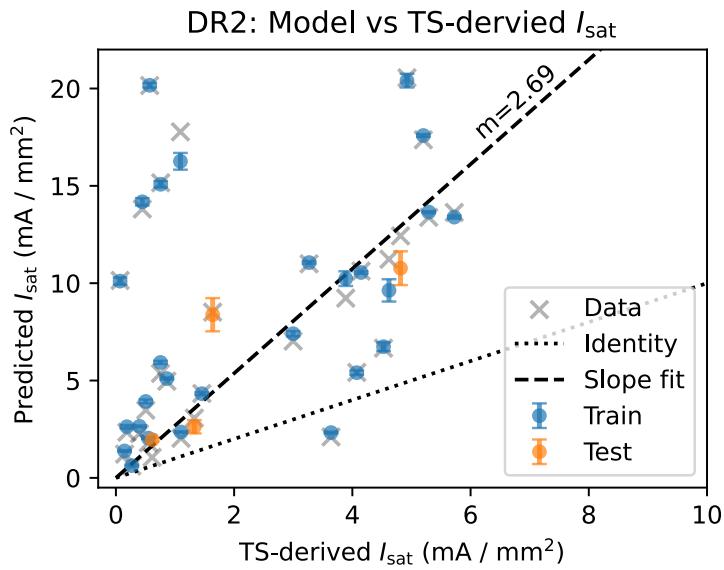
Figure 4.15: Comparison of original DR2 profiles with the profiles from the optimization dataset (DROpt) for the same machine configuration. The I_{sat} values in the DROpt dataset are not calibrated in this plot, indicating significant variation in probe calibration in this DROpt dataset.

4.6 Effect of I_{sat} calibration

The Langmuir probes did not seem to be behaving correctly when the optimization validation data were taken. The probes showed an *increasing* I_{sat} profile when moving further from the cathode in the lowest gas puff condition, which is in direct disagreement with previous measurements and intuition. An example of this discrepancy can be seen in Fig. 4.15, where a run from the original testing set (specifically DR2 run 10) is duplicated. The probes for the validation run can be either corrected for by assuming the 5 ms gas puff run has a flat axial profile, or normalizing the probes to the DR2 run 10 axial profile. Calibrating the probes using the DR2 run 10 reference was the best way to go because it corrects for both probe discrepancies as well as changes in the condition (or emissivity) of the main cathode.



(a) Thomson scattering (TS) 8 ms into the discharge compared to the model predictions (10 to 20 ms averaged). Broadly speaking, the TS measurement roughly agrees with the model estimate on average.



(b) Thomson scattering (TS) 12 ms into the discharge compared to the model predictions (10 to 20 ms averaged) and I_{sat} measurements one port away. The TS underestimates I_{sat} in general.

Figure 4.16

4.7 Inferring trends

A systematic study of the impact of discharge voltages on I_{sat} profiles has not been conducted using conventional techniques. Collecting both z- and x-axis profiles over a wide range of discharge voltages would take a considerable amount of time, mostly from the requirement to dismount and reattach the probes and probe drives along the length of the LAPD. This study has now been performed using the learned model, circumventing these time-consuming challenges. Model input parameters were chosen to be common, reasonable values: 1 kG flat field, 80 V gas puff, 38 ms gas puff duration, run set=DR2, and top gas puff off. The 38 ms puff is used in these predictions because it is the most common gas puff duration in the training set – the model is biased in favor of this gas puff setting. The results of changing the discharge voltage only can be seen in fig 4.17. Notably, the I_{sat} increases across both axes. Steeper axial gradients are seen with lower discharge voltages, but peaked x-profiles are seen at higher discharge voltages. The area closer to the source region (+z direction) appears to have a steep drop but flatter profiles down the length of the machine.

Unfortunately the discharge current was not included as an output in the training set. Otherwise the effect of changes in discharge power, rather than simply voltage, could be computed. The discharge current – and thus discharge power – is set by cathode condition, cathode heater settings, and the downstream machine configuration, and thus cannot be set to a desired value easily before the discharge. Discharge voltage, however, can remain fixed.

Of particular interest for some LAPD users is achieving the most uniform axial profile possible. We explore this problem in the context of mirrors. The gas puff duration is known to be a large actuator for controlling density and temperature and so is explored as a way of shaping the axial profile. We predict discharges with a flat 1 kG field with the probe in the center. The discharge voltage was set at 110 (a reasonable, middle value) with the run set flag=DR2 and top gas puffing=off. The inferred effect of gas puff duration on the axial gradient and axial gradient scale length can be seen in Fig. 4.18. Care was taken to handle the aleatoric (independent) uncertainty separately from the axially-dependent epistemic uncertainty. As seen in the figure, the mean

Discharge voltage: effect on x and z profiles

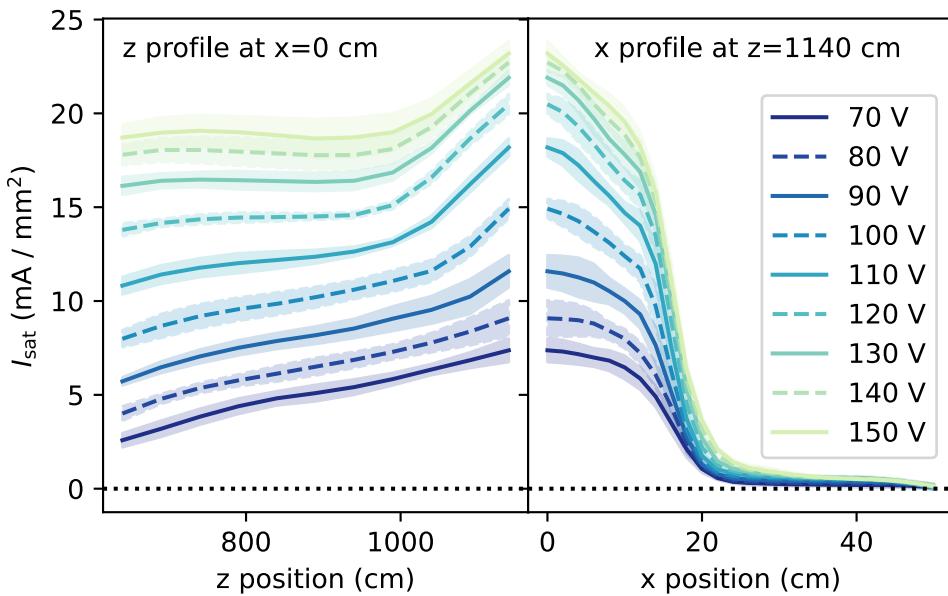


Figure 4.17: The z profile at $x=0$ cm and x profile at $z=1140$ cm for different discharge voltages.

The I_{sat} decreases with increasing voltage, and the error (filled regions) stays roughly the same, but in general increase slightly towards the cathode and at higher discharge voltages.

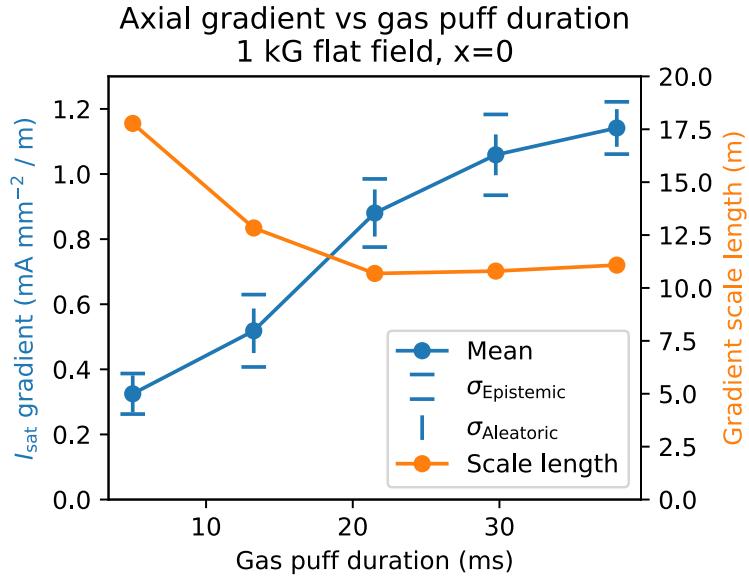


Figure 4.18: ML prediction: mean axial gradients decrease with decreased gas puff duration. Five durations are plotted between 5 and 38 ms (which are the bounds of the training data), evenly spaced. The gradient scale length also increases, indicating that the gradient change was not just from a decrease in the bulk I_{sat} .

axial gradient decreases when the gas puff duration is shortened. The gradient scale length also increases, so the mean gradient is not decreasing simply because the bulk I_{sat} is decreasing. This result suggests that the gas puff duration may be a useful actuator to consider when planning future experiments.

These applicability of these results are somewhat muted because the gas puff duration was not chosen randomly in the training discharges. Given this lack of data diversity, the accuracy and applicability of this study must be interpreted cautiously. When a model is trained on *all* data available (using the cross-validated test set MSE as a guide for error), which includes the 20 ms gas puff case, the predicted gradient scale length decreases uniformly across the duration scan by 1 meter. The fact that the trend remains intact when an additional, randomized intermediate gas puff case is added gives confidence in the predictions of the model despite the lack of data diversity.

4.8 Optimizing profiles

One particular issue seen in LAPD plasmas is sharp axial density and temperature gradients. Some experiments require flat gradients, such as Alfvén wave propagation studies. We explore optimizing the axial I_{sat} variation as an approximation to this problem. In addition, in this case the optimization problem is used as a tool to evaluate the quality of the learned model. This is a very demanding task because the trends inferred by the model along all inputs must simultaneously be accurate. Constraints on this optimization further increase the difficulty of the problem. Success in optimization provides strong evidence that the model has inferred relevant trends in predicting I_{sat} . We quantify the uniformity of the axial profile by taking the standard deviation of I_{sat} of 11 points along the z-axis ($x, y = 0$). The required LAPD state for attaining the most uniform axial profile can be found by finding the minimum of this standard deviation with respect to the LAPD control parameters and flags:

$$\text{Inputs} = \arg \min_{\substack{\text{Inputs} \neq z}} \text{sd}(I_{\text{sat}}|_{x=0}) \quad (4.3)$$

The largest axial variation can likewise be found by finding the maximum. The model inputs used for this optimization can be found in Table 4.2.

For this optimization we use an ensemble of five β -NLL-loss models with weight decay $\lambda = 0$. The $\lambda = 0$ model is used because it appears to give the most useful uncertainty estimate (seen in Fig. 4.12). The optimal machine actuator states are found by feeding a grid of inputs into the neural network. This variance estimate is not well-calibrated: the error of the predictions on the test set falls far outside the predicted uncertainty. However, this uncertainty can be used in a relative way: when the model is predicting far outside its training range, the predicted variance is much larger. The ranges of inputs into this model are seen in Table 4.2. These inputs yield 127,234,800 different machine states (times five models) which takes 151 seconds to process on an RTX 3090 (≈ 4.2 million forward passes per second) when implemented in a naive way. The number of forward passes can be reduced by a factor of 51 if the x value is set to 0 cm. Note that gradient-based methods can be used for search because the network is differentiable everywhere but this

Table 4.2: Machine inputs and actuators for model inference

Input or actuator	Range	Step	Count
Source field	500 G to 2000 G	250 G	7
Mirror field	250 G to 1500 G	250 G	6
Midplane field	250 G to 1500 G	250 G	6
Gas puff voltage	70 V to 90 V	5 V	5
Discharge voltage	70 V to 150 V	10 V	9
Gas puff duration	5 ms to 38 ms	8.25 ms	5
Probe x positions	-50 cm to 50 cm	2 cm	51
Probe y positions	0 cm	—	—
Probe z positions	640 cm to 1140 cm	50 cm	11
Probe angle	0 rad	—	—
Run set flag	off and on	1	2
Top gas puff flag	off and on	1	2

network and parameter space is sufficiently small that a comprehensive search is computationally tractable.

Like any optimization method, the results may be pathologically optimal. In this scenario, the unconstrained minimal axial variation is found when the I_{sat} is only around 1 mA/mm², which is quite small and corresponds to $1-2 \times 10^{12}$ cm⁻³ depending on Te. The inputs corresponding to this optimum are given in the second column of Table 4.3. This density range is below what is required or useful for many studies in the LAPD.

Since many physics studies require higher densities, we constrain the mean axial I_{sat} value to be greater than 7.5 mA/mm² (roughly $0.5-2 \times 10^{13}$ cm⁻³). The “run set flag” is set to “on” for cases to be validated (bolded in Table 4.3) because we wish to keep the turbopumps on to represent typical LAPD operating conditions. In addition the “top gas puff flag” was set to ‘off’ to minimize the complexity of operating the fueling system on followup dataruns and experiments. Turning the top gas puff valve on is predicted to decrease the average I_{sat} by -2 mA/mm² for strongly varying profiles, but otherwise the shapes are very similar. The inputs corresponding to the maximum and minimum axial variation under these constraints can be seen in columns 3 and 4 of Table 4.3. For contrast we also consider what machine settings would lead to the greatest axial variation. The results of both of these optimizations can be seen in Fig. 4.19. The optimizations yield profiles that have the largest I_{sat} values closest to the cathode, which is expected.

The prediction for an intermediate axial variation case is also seen in Fig. 4.19. The intermediate case was chosen as somewhere around half way between the strongest and weakest case with a round index number (15000, in this case). The parameters for intermediate case are also enumerated in Table 4.3.

The predicted configurations with the run set flag on and top gas puff flag off (bolded in Table 4.3 were then applied on the LAPD. The data collected, compared with the predictions can be seen in Fig. 4.19.

For the optimized axial profiles, the absolute value of the I_{sat} predictions compared to measure-

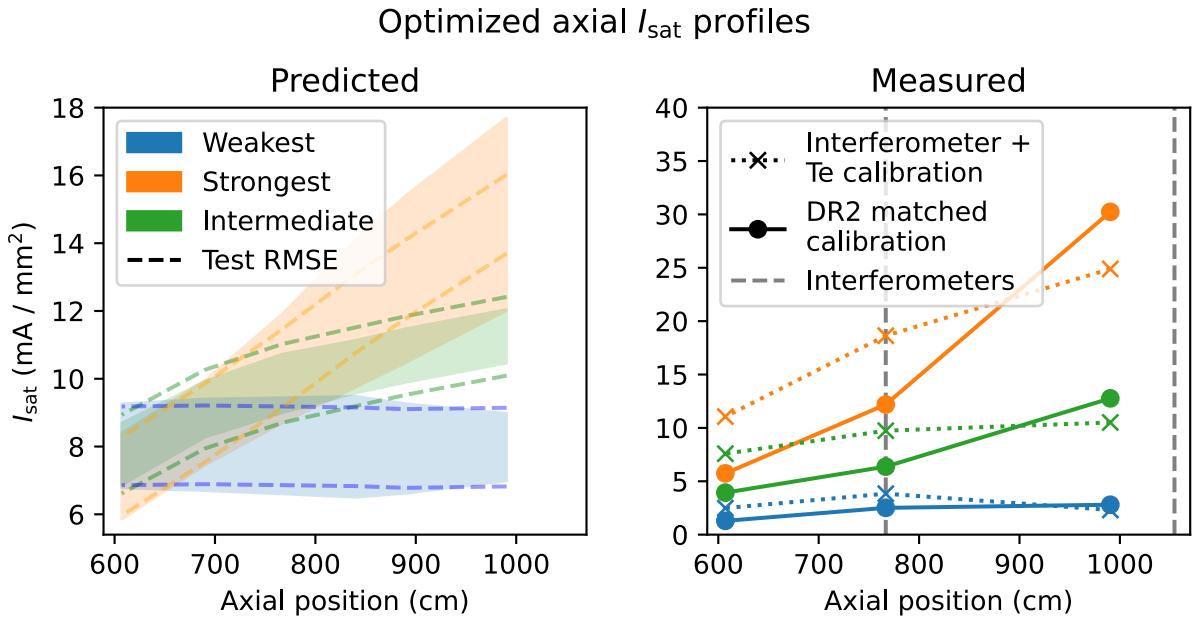


Figure 4.19: Axial profiles, predicted and measured, for the optimized weakest (blue), intermediate (green), and strongest (orange) cases. a. The shaded region covers the mean prediction \pm one standard deviation, and the dashed lines are \pm the median cross-validation RMSE values. b. The measured I_{sat} values are calibrated to DR2 run 10 (solid lines), or using triple probe Te measurements on the probe and linearly extrapolating the interferometer measurements (dotted lines). The absolute values disagree between the predicted and measured values, but axial trends are consistent with the optimization.

Table 4.3: Machine inputs and actuators for optimized axial profiles

Input or actuator	Weakest	Weakest	Strongest	Intermediate
I_{sat} constraint (mA/mm ²)	$I_{\text{sat}} = \text{any}$	$I_{\text{sat}} > 7.5$	$I_{\text{sat}} > 7.5$	$I_{\text{sat}} > 7.5$
Source field	750 G	1000 G	500 G	2000 G
Mirror field	1000 G	750 G	500 G	1250 G
Midplane field	250 G	250 G	1500 G	750 G
Gas puff voltage	70 V	75 V	90 V	90 V
Discharge voltage	130 V	150 V	150 V	120 V
Gas puff duration	5 ms	5 ms	38 ms	38 ms
Run set flag	on	on	on	on
Top gas puff flag	on	off	off	off

ment do not agree. All of the predicted profiles have overlapping predictions (within the predicted error) at the region furthest from the cathode, but the measured values do not show that behavior. Although the mean I_{sat} value was constrained to be greater than 7.5 mA/mm², the measured mean was 2 mA/mm² for the weakest case.

The important result is that the optimized LAPD settings, when implemented on the machine, do yield profiles with strong, intermediate, and weak axial variation. Although the minimum- I_{sat} constraint was violated for the case of weakest axial variation case, this optimization would nonetheless be very useful for creating axial profiles of the desired shape.

There are three contributing factors to the mismatch of the ML-predicted values and the real measured values. First, the condition of the machine, such as the cathode emissivity or temperature or the downstream neutral pressure, are unquantified and cannot be compensated for in data preprocessing or in the model itself. Second, the calibration of the Langmuir probes could differ substantially between runs. The probes in the training data run sets (DR1 and DR2) were well-calibrated to each other within the run set, but were not absolutely calibrated. The probes used for verifying the optimization were not calibrated. A rough calibration was performed by linearly

extrapolating interferometer measurements and using triple probes (dotted lines on the right panel in Fig. 4.19). A configuration identical to DR2 run 10 was also measured to simultaneously correct cathode condition and probe calibration (solid lines on the right panel in Fig. 4.19). Langmuir probe calibration is discussed further in Appendix 4.6. Third, the original dataset may not have sufficient diversity to make accurate predictions on such a constrained optimization problem.

If this optimization were performed using the dataset instead of the model, the constrained search would encompass just 10084 shots out of the 131550 shots total in the training dataset, or around 7.7%. Including the on-axis constraint reduces the number of shots down to 303 (270 in the training set), or 0.23% of all shots in the dataset. We conclude that this optimization of an arbitrary objective function, as done here, would be intractable using traditional, non-machine learning techniques because orders of magnitude more dataruns would need to be collected.

Optimization requires correctly learning the trends of all inputs and how they interact. In addition, as seen from the shot statistics, the model was trained on very few shots in the constrained input and output space. These two factors – the need for the model to learn all trends and the constrained search space – combine to make an incredibly difficult task that functions as a benchmark for the model. These factors considered, it is not surprising that the model incorrectly predicts the absolute value. The uncertainty predicted by the model, though not well-calibrated, was nonetheless very large compared to the median test set RMSE. The model did predict the trends correctly, however; the optimized, measured profiles were strong, intermediate, or weak.

We did not check to see if the predicted optima were actually true optima: an approximation of the local derivative using a finite-difference technique would require much more run time on the LAPD than was available.

4.9 Discussion

4.9.1 Key achievements

To the authors' knowledge, this work is one of the first instances machine learning has been used to infer specific trends and optimize profiles in magnetized plasmas. Three examples of trend inference were shown in this paper: influence of magnetic geometry on plasma width, changes in the axial and radial profiles with changing discharge voltage, and the relationship of gas puff duration with axial gradient scale length. In addition, the axial profile was optimized by minimizing (or maximizing) the axial standard deviation. There is no other way of simultaneously uncovering many trends or finding optima without using an ML model trained on a diverse dataset. Traditionally, such studies would require extensive scans over grids to map the parameter space, but here it was accomplished with a relatively small amount of data.

The trends inferred in this work, such as changing discharge voltages, gas puff durations, or mirror fields, would traditionally require a grid scan (varying one parameter at a time) in LAPD settings space. Here instead we are able to extract any trend covered by the training set with only a minimal amount of machine configurations sampled. Both data collection runs lacked absolute I_{sat} calibration and had potential differences in cathode condition. Despite these issues the model learned trends that were exploited via optimization.

In addition, this work demonstrates uncertainty quantification broken down into epistemic and aleatoric components by using ensembles and a negative-log likelihood loss. This uncertainty estimate is useful in gauging relative certainty between different predictions of the model which increases confidence in the predictions of the model. In general, the total uncertainty predicted by the model increases when predictions are made outside the bounds of the training data.

Fundamentally, this model can predict I_{sat} with uncertainty at any point in space covered by the training data. No other model currently exists that can perform this prediction. Traditionally, this capability would be possible only with a detailed theoretical study.

4.9.2 Current limitations

This study would be dramatically improved by collecting more, diverse data. Only 44 of the 67 dataruns in this dataset had randomly sampled LAPD machine settings which is very small compared to the over 60,000 possible combinations. In addition, there are many other settings or parameters that were not changed in this study, such as gas puff timings, gas puff valve asymmetries, wall/limiter biasing, cathode heater settings, operation of the north cathode, and so on. The bounds of the inputs were also conservative; all settings in this study could be pushed higher or lower with a small amount of risk to LAPD operations. In addition, the placement of the probes could be further varied and placed outside the mirror cell, which would provide a more complete picture of LAPD plasmas, particularly axial effects.

Probe calibrations differed between the two training run sets (DR1 and DR2) – and a flag was introduced to distinguish between them – but despite this deficiency combining the two run sets was shown to be advantageous for model performance. The condition of the cathode (e.g., electron emissivity and uniformity) also has a large impact on the measured I_{sat} . The improved model performance with the flag suggests that inconsistencies between dataruns could be compensated for using latent variables if a generative modeling approach is to be taken. At the very least, this model provides a way to benchmark these differences in machine state.

Concerning the model, hyperparameter tuning could be performed. In this study a few extra percent in MSE performance is not meaningful considering the limited dataset. Instead, we focused on the trends and insights that can be extracted from this model rather than simple predictive accuracy. There may also be regimes in hyperparameter space where the uncertainty is better calibrated (perhaps using early stopping). Uncertainty estimation is important, even if the absolute uncertainty is not well-calibrated because it can provide a useful relative estimate as shown in this paper.

Trends such as the dependence of axial gradient on the gas puff duration (fig. 4.18) or the effect discharge voltage on x-z profiles (fig. 4.17), although intuitive, remain unverified. Verification of

these trends would increase confidence of model predictions when setting LAPD parameters in future experiments.

4.9.3 Future directions

The neural network architecture used here can readily scale to additional inputs and outputs; including time-series signals is the obvious next step. Integration of multiple diagnostics – perhaps starting with individual models before combining them – could enable inference of plasma parameters throughout the device volume. For example, combining triple probe electron temperature measurements with existing I_{sat} data would allow density predictions anywhere in the plasma. This capability could enable in-situ diagnostic cross-calibration (e.g., the Thomson scattering density measurement) and prediction of higher-order distribution moments like particle flux. The model could be further enhanced by incorporating physics constraints such as boundary conditions (e.g., zero I_{sat} at the machine wall) or symmetries.

The problem presented here – learning time-averaged I_{sat} trends – is fairly simple and required a relatively simple model. As demonstrated in this work, ML provides a way to explore regions of parameter space quickly and efficiently. Most physics studies on plasma devices (and fusion devices) are dedicated to a single particular problem, use grid scans, and are not useful to other experiments or campaigns. This work shows a way of using data and trends uncovered from other experimental studies. This work also demonstrates that random exploration can be a useful tool: the increased diversity of the aggregated data will generally benefit an ML model whether the experimenter discovers something new or not.

4.10 Conclusion

We demonstrate the first randomized experiments in a magnetized plasma experiment to train a neural network model. This learned model was then used to infer trends when changing field configuration, discharge voltage, or gas puff duration. This model was also used to optimize axial

variation of I_{sat} as measured by the standard deviation, which was validated in later experiments despite poor absolute error.

We strongly advocate that all ML-based analyses in plasma and fusion research should be validated and used to gain insight by inferring trends, as demonstrated here. This validation step is crucial for ensuring that ML models capture physically meaningful relationships and the insights provided may provide direction for future research. We hope this is the first step towards automating plasma science.

4.11 The open dataset and repository

All the code to perform the ML portion of this study is available at <https://doi.org/10.5281/zenodo.1500783>. The training datasets are also available in that repository in the datasets directory. Additional data are available upon request. The repository also contains additional training details and the notebooks for generating the plots in this document. The plots used in this paper were made in jupyter notebooks, which are also uploaded. The final training code can be found in `train_dense_beta_NLL.py`. Trained models are found in the `code/training_runs` directory. The history of many training runs can be found on Weights and Biases: <https://wandb.ai/phil/profile-predict> and the accompanying notes on these trained models are found in the associated pdf on github. The code and dataset are licensed under Creative Commons Attribution Share Alike 4.0 International.

CHAPTER 5

Reconstructing missing diagnostics using energy-based models

The goal here is to show that diagnostics can be reconstructed from other diagnostics and machine state information (MSI) using energy-based models (EBMs). This is the first time that EBMs have been used and possibly the first time diagnostics reconstruction has been attempted using neural networks in a magnetized plasma context. The bulk of this work was performed in 2021 and 2022.

5.1 Brief introduction to energy-based models (EBMs)

5.2 Building a 15 million shot dataset

The LAPD has a set of permanent diagnostics and machine state information (MSI) that is produced but was not recorded for every shot. I built out a diagnostics pipeline that would record the following for every single shot, continuously:

1. Discharge current (time series)
2. Discharge voltage (time series)
3. 56 GHz interferometer (time series)
4. Five axially spaced photodiodes, one having a He-II filter (time series)
5. RGA and total gas pressures (single point for each mass)
6. Magnetic field profile

These data were recorded so that the shots could be synchronized with the data acquisition system used for probe measurements. A plot of a couple of shots can be seen in fig. ??.

These diagnostics and MSI have some particular characteristics. The discharge current may not be accurate based on how the value is measured. The interferometer skips when the plasma density gets too high and the signal cannot be recovered. However, the interferometer signal is still included because these fringe skips (and timing thereof) contain information about the plasma density evolution, though the numerical value may be difficult to acquire. The diodes are not absolutely calibrated, have a nonlinear response, and are not restricted to the visible spectrum. The RGA takes a minute or two to sweep, so partial pressures are not necessarily exact for the particular shot it is recorded. The magnetic field profile is calculated, not measured, so some field measurements may be incorrect if the state of a magnet power supply is recorded incorrectly, or if the source field currents were not manually updated. Despite all these issues, we proceed.

TODO: Plot of example data

This system recorded shots continuously, capturing a wide variety of LAPD machine conditions and experiments. Some particular challenges with this dataset was that other important information was not recorded in an orderly way. The north LaB6 source current and voltage time traces were not recorded even though the north source parameters could have a significant impact on the behavior of the plasma. The locations of large antennas or obstructions were also not recorded conveniently: retrieving this information would require manually opening and reading many experimental datarun files, but that only corresponds to probe data. In general this information was not recorded between dataruns.

In addition, the data may be highly biased towards a particular set of cases. The LAPD was often left as-is without changing parameters for considerable lengths of time, either between dataruns or campaigns, or in general for a particular experiment week. Thus, the data collected may not contain sufficient diversity to learn trends well, particularly with important machine state information not recorded.

The general cathode (and thus plasma condition) was also left unquantified. Variation emissivity across the cathode could lead to considerable changes across the plasma column, and the discharge characteristics could drift over time. This unquantified drift in plasma conditions further makes this dataset a difficult training target, given that the same machine configuration could lead to multiple types of plasma.

TODO: show plot of long-term trends of the discharge current and density changing.

5.3 Training the model

5.4 Unconditional sampling

5.5 Reconstructing missing diagnostics via conditional sampling

5.6 Inferring trends and outlook

CHAPTER 6

Conclusions

6.1 Future Directions

APPENDIX A

0D mirror optimization

1. **problem:** Main things that need to be done: find references and double check the equations
 2. **problem:** Power to central cell isn't accounted for in plug temperature caluations fudge factor is used
-
1. **TODO:** Check that DT alpha orbit is contained
 2. **TODO:** Add HHFW heating to increase Einj above Ebeam
 3. **TODO:** Add neutron dpa
 4. **TODO:** Implement assumption calculations for FLR effects and the paraxial approximation
 5. **TODO:** Add FLR stabilization estimate (Eq 38 in "Magneto-hydrodynamically stable[...]", Ryutov 2011)
 6. **TODO:** Add neutral beam shine through as a condition for plasma density
 7. **TODO:** Compare to baseline in section 7 of Egedal et al 2022 [EEF22]
 8. **TODO:** Implement beta-enhanced mirror ratio limits from diamagnetic-bubble paper. Beta-enhanced mirror ratio flag?
 9. **TODO:** Add "tail wagging" stabilization power cost
 10. **TODO:** Calculate stability thresholds and growth rates

A.1 List of assumptions / conditions

There are many issues and assumptions with this analysis (in no particular order):

1. Powers and particles are not strictly balanced in tandem mirrors
2. Thermal barriers are ignored which may be very important for a practical reactor
3. A fudge factor is used for electron temperatures when plug electrons are heating the central cell
4. Macrostability is not considered
5. Microstability is not considered
6. Plug confinement time is not self-consistent with plug temperatures
7. Effects of field ripple are not calculated
8. T-T and T-He3 reaction rates are not considered
9. Radial transport is assumed to be only classical
10. Cat-DD assumes instant burnup of products which is unreasonable, particularly at the high ion energies needed in mirror reactors
11. Impurities are assumed to be zero
12. Heating and magnet costs are not justified
13. All fusion power exits the plasma immediately (charged particles are collected by the direct-energy converter, neutrons absorbed by the blanket)
14. When using the temperature model from Egedal 2022 [EEF22], we assume that the auxiliary power is much less than the beam power ($P_{aux}/P_{NBI} \ll 1$) or else the model may be inaccurate. Auxiliary power (say, to compensate for classical diffusion losses or additional ECH) can be included in this model but it would require iterative solving.

15. Burnup fraction is sufficiently small that fusion reactions are not a significant loss of fuel (ironically).
16. The DECs collect all ion losses at a fixed efficiency

A.2 User specified parameters

A.2.1 Simple mirror endplug

1. Mirror field, plug (T): $B_{p,m}$
2. Plug (i.e., midplane) cell field (T): B_p
3. Magnet bore/throat radius (m): r_b
4. Plug length (m): L_p
5. Neutral beam energy (keV): E_{inj} or E_b
6. Beta limit (critical stability): β_{limit} (set to 0.8 [KLZ21])
7. Effective species mass (amu): μ
8. Effective atomic number (account for He and other impurities): Z_{eff}

The β_{limit} (discussed in Kotelnikov 2021 [KLZ21]) assumes a stationary plasma (no rotation, no flow out the ends), ignores finite-Larmor-radius (FLR) effects (which stabilize $m > a^2/L\rho_i$ modes), and uses the paraxial approximation ($L_m \gg a$). It also assumes $\beta \ll 1$ but this paper shows that these results match up with GDT experiments. The β_{limit} depends on the radial pressure profile falloff; the faster the falloff, the lower the β_{limit} . L is the length from midplane to throat, and L_m is the length of the mirror (highest field to lowest, I think). Profile calculations will not be included in a 0D optimization. The relevant assumptions for FLR effects and the paraxial approximation should be calculated and shown in the output to make sure they are not dramatically violated.

A.2.2 Tandem mirror

Central cell parameters defined below. Simple mirror endplugs are used on either end.

1. Central cell field (T): B_{cc}
2. Central cell to plug density ratio: n_{cc}/n_p
3. Central cell ion to plug electron temperature ratio: $T_{cc,i}/T_{p,e}$ (assumes Maxwellian)
4. Central cell to plug electron temperature ratio: $T_{cc,e}/T_{p,e}$
5. Central cell length (m): L_{cc}
6. Electron temperature fudge factor: $T_{ep, \text{fudge}}$ if electron are heating the central cell. Default value is 0.5

A.2.3 Engineering parameters

1. Vessel wall radius (with respect to plasma radius): $a_{\text{wall}} = a_{\text{wall, ratio}} \cdot a_{\text{plasma}}$
2. Blanket thickness: d_{blanket}
3. Vacuum vessel thickness: d_{vv}
4. Direct Converter Efficiency (used in the mirror exhaust): η_{DEC}
5. Thermal to electric conversion efficiency: η_{TE}
6. ECH heating efficiency: η_{ECH}
7. NBI heating efficiency: η_{NBI}
8. RF heating efficiency: η_{RF}

Optimizing the blanket and vacuum vessel thickness would probably require some neutronics calculations which would probably depend on the fuel mix, so we're just going to leave those constant in our optimization.

A.3 Fusion

DT fusion helium energy (keV): $E_\alpha = 3500$

A.3.1 Reactivity

DD and DT fusion cross-section parameterizations can be found in Bosch 1992. [BH92]. What we care most about is the fusion reaction rate per unit volume (eq. 10 in the paper):

$$\frac{dR}{dV} = \frac{n_i n_j}{1 + \delta_{ij}} \langle \sigma v \rangle \quad (\text{A.1})$$

This parameterization accepts ion temperature in keV and gives reactivity in cm^3/s :

$$\langle \sigma v \rangle = C1 \cdot \theta \sqrt{\xi / (m_r c^2 T^3)} e^{-3\xi} \quad (\text{A.2})$$

$$\theta = T / \left[1 - \frac{T(C2 + T(C4 + TC6))}{1 + T(C3 + T(C5 + TC7))} \right] \quad (\text{A.3})$$

$$\xi = (B_G^2 / 4\theta)^{1/3} \quad (\text{A.4})$$

$$B_G = \pi \alpha Z_1 Z_2 \sqrt{2 m_r c^2} \quad (\text{A.5})$$

where m_r is the reduced mass and α is the fine structure constant. The coefficients ($C1$, $C2$, and so on) are in the paper cited above. This parameterization is valid for T_i between 0.2 to 100 keV.

problem: We will definitely exceed this in our optimization – the extent of the deviation should be quantified. The cross section explodes for DD after roughly 500 keV Max error is 0.25% for DT and 0.35% and 0.3% for DD $\Rightarrow p$ T and DD $\Rightarrow n$ He3, respectively. phil: We will want to use the cross-section (the parameterization of which goes up to 550 keV for DT in [BH92]) and integrate over ion distribution. Integration over non-Maxwellian ion distributions should be doable if FBIS [EEF22] can give the full distribution. phil: For now I'm just linearly interpolating (in

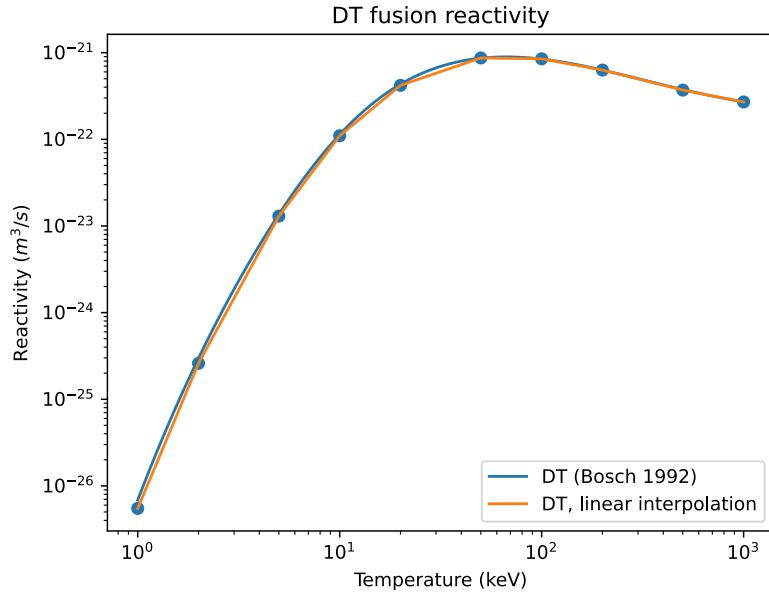


Figure A.1: DT reactivities

log-log space) between the points provided in the NRL. Comparison plots of the different methods for DD and DT can be seen in figures A.1 and A.2.

A.3.2 Fusion power

DT fusion reaction rate (#/s):

$$R_{x,DT} = V n_D n_T \langle \sigma v \rangle_{DT} \quad (\text{A.6})$$

If $n_D = n_T = n/2$, then this becomes $V \frac{n^2}{4} \langle \sigma v \rangle_{DT}$

DD fusion reaction rate (#/s):

$$R_{x,DD} = V \frac{n_{\text{plug, D}}^2}{2} \langle \sigma v \rangle_{DD} \quad (\text{A.7})$$

The $\frac{1}{2}$ factor is to avoid double counting DD reactions.

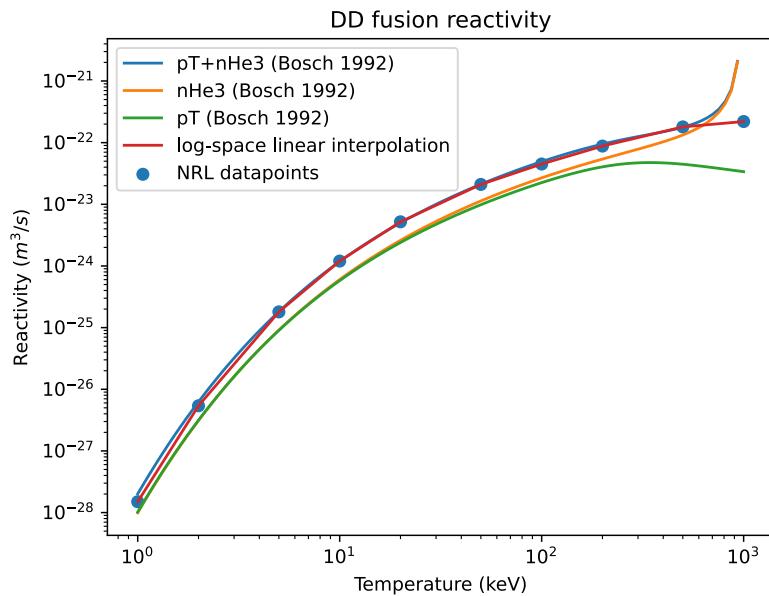


Figure A.2: DD reactivities

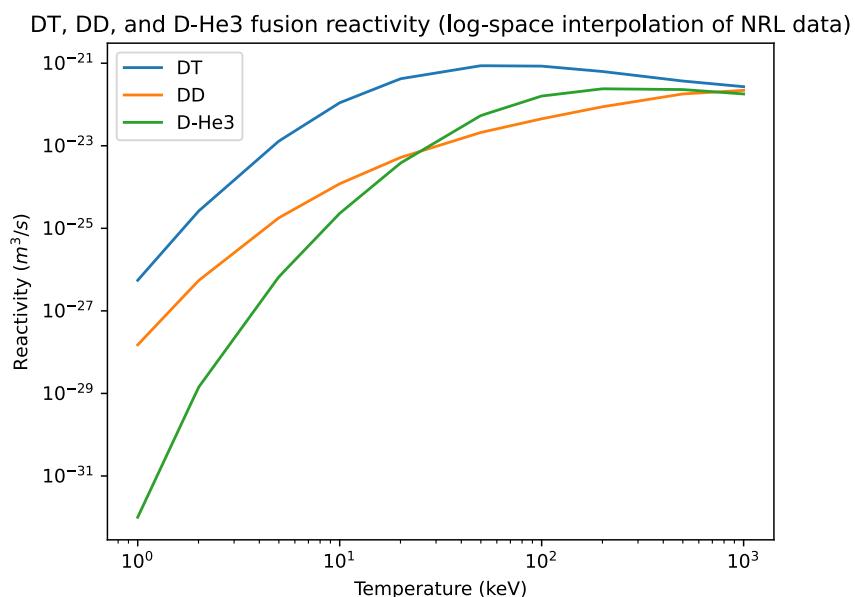


Figure A.3: DT, DD, and D-He3 reactivity comparison

Fusion power (MW):

$$P_{DT,n} = 14.1 |e| R_{x,DT} \quad (A.8)$$

$$P_{DT,+} = 3.5 |e| R_{x,DT} \quad (A.9)$$

$$P_{DD,n} = 2.45 |e| R_{x,DD} \cdot \frac{1}{2} \quad (A.10)$$

$$P_{DD,+} = (4.02 + 0.82) |e| R_{x,DD} \cdot \frac{1}{2} \quad (A.11)$$

$$(A.12)$$

It's useful to split the power into charged and neutrons because energy is extract from them in different ways. Neutrons provide thermal power, charged particles heat the plasma and/or are directly captured by the DECs. The $\frac{1}{2}$ coefficient on the DD reactions assumes a 50-50 split on the DD branching ratio which actually varies with energy and may be significant above around 100 keV. If we assume the tritium produced from a DD reaction is burned instantly, then the additional power produced ("catalyzed DD") is:

$$P_{cat\ DD,n} = 14.1 |e| R_{x,DD} \cdot \frac{1}{2} \quad (A.13)$$

$$P_{cat\ DD,+} = (3.5 + 18.3) |e| R_{x,DD} \cdot \frac{1}{2} \quad (A.14)$$

$$(A.15)$$

We assume the tritium is burned instantly because the DT reaction rate is much higher than DD and D-He3 fusion up to around 200 keV, after which it's only slightly higher up to around 1 MeV. A more accurate estimate of fusion power would require estimates of D-He3, TT, and T-He reaction rates and density evolution of each species. A plot of reactivities can be found in Fig. A.3. phil:
I do not know how to do this steady-state calculation in a simple 0d way – perhaps this sort of calculation is left for a more detailed reactor model.

A.4 General formulae

Electron cyclotron frequency (GHz):

$$f_{\text{ECH}} = \frac{eB}{2\pi m_e c} = 28B \quad (\text{A.16})$$

Ion cyclotron frequency (MHz):

$$f_{ci} = \frac{ZeB}{2\pi m_i c} \quad (\text{A.17})$$

$$f_{ci,D} = 7.63B \quad (\text{A.18})$$

$$f_{ci,T} = 5.09B \quad (\text{A.19})$$

Here, Z is the charge state of the ion.

Electron plasma frequency (Hz):

$$f_{pe} = \frac{1}{2\pi} \sqrt{\frac{4\pi n_e e^2}{m_e}} = 8.98 \cdot 10^3 \sqrt{n_{\text{plug}}} \quad (\text{A.20})$$

Ion plasma frequency (Hz):

$$f_{pi} = \frac{1}{2\pi} \sqrt{\frac{4\pi n_i Z^2 e^2}{\mu m_p}} \quad (\text{A.21})$$

$$f_{pi,D} = 2100 \sqrt{\frac{n_{\text{plug}}}{2}} \quad (\text{A.22})$$

$$f_{pi,T} = 2100 \sqrt{\frac{n_{\text{plug}}}{3}} \quad (\text{A.23})$$

Here, μ is the mass of the ion in proton mass units (e.g. $\mu_{\text{Deuterium}} = 2$ and $\mu_{\text{Tritium}} = 3$).

Lorentz factor (γ):

$$\gamma = \sqrt{1 + \frac{T_e}{m_e c^2}} = \sqrt{1 + \frac{T_e}{511 \text{keV}}} \quad (\text{A.24})$$

Ion thermal velocity:

$$v_{Ti} = 97900 \sqrt{\frac{10^3 E_{\text{ion}}}{\mu}} \quad (\text{A.25})$$

Ion gyroradius: problem: Cary's spreadsheet uses the temperature to calculate this but calls it " E_{ion} ". I think temperature should be used to accurately account for the energy in vperp

$$\rho_i = \frac{mv_\perp}{qB} = 3.22 \cdot 10^{-3} \frac{\sqrt{\mu E_{ion}}}{B_p} \quad (\text{A.26})$$

Whistler wavelength:

$$\lambda_{\text{whistler}} = \sqrt{\frac{2\pi\Omega_e c^2}{\Pi_e^2 f}} \quad (\text{A.27})$$

$$\lambda_{\text{whistler}} = \sqrt{\frac{90 f_{\text{ECH}}}{f_{pe}^2 f_{D,2\text{nd Harmonic}}}} \quad (\text{A.28})$$

The 2nd formula is what appears on the spreadsheet and is used to estimate the size of the RF wave used for HFW as compared to the size of the plasma. It takes into account the various constants and units used in the spreadsheet.

Collision rates (from NRL):

$$v_e = 2.91 \cdot 10^{-6} \frac{n_e \ln \Lambda}{T_e^{3/2}} \quad (\text{A.29})$$

$$v_i = 4.80 \cdot 10^{-8} \frac{Z_{\text{eff}}^4 n_i \ln \Lambda}{\mu^{1/2} T_i^{3/2}} \quad (\text{A.30})$$

These can be rearranged to give the following collision times (s): problem: this seems to assume $\log \Lambda = 20$ (should be doubled check but doesn't probably make much of a difference – comparing to eq 4.8 in [Bal77]). problem: Depends on Zeff which isn't accounted for during optimization/power balance. phil: Technically we should use $T = \frac{2}{3}E_{ion}$ for the central cell since the distribution is likely approximately Maxwellian, but it only changes τ_{ii} by a factor of 1.8.

$$\tau_{ee} = 10^{-4} \frac{T_e^{3/2}}{n_{20} \lambda_{ee}} \quad (\text{A.31})$$

$$\tau_{ii} = 1.25 \cdot 10^{-4} \frac{\mu^{1/2} E_{ion}^{3/2}}{n_{20} Z_{\text{eff}}^4} \quad (\text{A.32})$$

Slowing down times [Dol82] phil: I assume Z is Zeff? Need to double check. I put it as Zeff in the ipynb :

$$\tau_{i,\text{slow}} = 0.1 \frac{\mu T_e^{3/2}}{n_{20} Z^2 \lambda_{ei}} \quad (\text{A.33})$$

$$\tau_{i,\text{fast}} = \left(\left(\tau_{ii} 0.4 \log \frac{R_m}{\sqrt{1-\beta}} \right)^{-1} + \frac{1}{\tau_{\text{slow}}} \right)^{-1} \quad (\text{A.34})$$

Here, we can write $Z = 2$ for alpha particles. The 2nd equation comes from substituting the expression for T_e in a purely NBI heated case seen above.

Coulomb logarithms:

$$\lambda_{ee} = 23.5 - 0.5 \ln n_e + 1.25 \ln T_e - \left(10^{-5} + \frac{(\ln T_e - 2)^2}{16} \right)^{1/2} \quad (\text{A.35})$$

$$\lambda_{ei} = 24 - 0.5 \ln n_e + \ln T_e \quad (\text{A.36})$$

$$\lambda_{ii,\text{Cary}} = 31 - 0.5 \ln n_e + \ln T_e \quad (\text{A.37})$$

$$\lambda_{ii,\text{NRL}} = 23 - 0.5 \ln n_e + 1.5 \ln T_i \quad (\text{A.38})$$

The formula for λ_{ee} is from NRL. The formula for λ_{ei} is from NRL. However, the plasma does not fit into any of the 3 limiting cases described in the formulary. We have picked the formula based on the condition that is violated the least severely. There are 2 formulas for λ_{ii} . They do not have a large disagreement in the ranges of T_e and T_i of interest.

A.5 Radial particle transport

As of the time of writing, diffusive radial transport in mirror reactors appears to be an open question. The goal here is to provide reasonable estimates of radial particle loss and how each scale, not necessarily going for high-accuracy predictions (though being close would be nice!)

A.5.1 Classical diffusion

Assuming Fick's law and a linear density gradient from $3n$ to 0 (from Chen 5.8):

$$\tau_{\text{classical}} = \frac{nV}{A \cdot \Gamma} = \frac{na}{-2D_{\perp} \nabla n} \quad (\text{A.39})$$

where D_{\perp} is defined as

$$D_{\perp} = \eta_{\perp} n \sum T / B^2 \quad (\text{A.40})$$

and the parallel and perpendicular conductivities are

$$\eta_{||} = 5.2 \cdot 10^{-5} \frac{Z \ln \Lambda}{T^{3/2}} \quad (\text{A.41})$$

$$\eta_{\perp} = 2 \cdot \eta_{||} \quad (\text{A.42})$$

Combining all of this together gives

$$\tau_{\text{classical}} = \frac{a^2 B^2 T_e^{3/2}}{3.12 \cdot 10^{-4} \cdot n Z \sqrt{\mu} \ln \Lambda \sum T} \quad (\text{A.43})$$

A.5.2 Bohm diffusion

phil: I don't like how pessimistic the estimates are for this are. Reasonable values gets you 500 μs confinement times for Bohm diffusion. Bohm diffusivity:

$$D_{\text{Bohm}} = \frac{1}{16} \cdot \frac{T_{i,cc} \cdot 10^3}{B_{cc}} \quad (\text{A.44})$$

Normalized gyroradius (assuming deuterium):

$$r_{\text{Larmor}} = \frac{\sqrt{2mE_{\perp}}}{2eB_{cc}} = \frac{0.00791 \sqrt{T_{i,cc}}}{B_{cc}} \text{ cm} \quad (\text{A.45})$$

$$\rho^* = \frac{r_{\text{Larmor}}}{a_{cc}} \quad (\text{A.46})$$

Again, using Fick's law and assuming a linear density gradient from $3n$ to 0 (so that total particle number remains $n \cdot V$), cross-field particle flux is:

$$\Gamma = -D_{\text{Bohm}} \cdot \nabla n_e \quad (\text{A.47})$$

$$= \frac{1}{16} \cdot \frac{T_{i,cc} \cdot 10^3}{B_{cc}} \cdot 3n_i \quad (\text{A.48})$$

which implies a characteristic confinement time of

$$\tau_{\text{Bohm}} = N_{\text{tot}} \left/ \frac{dN}{dt} \right. = n_i \cdot V / (\Gamma * A) \quad (\text{A.49})$$

$$= n_i \cdot \pi a^2 L \left/ \left(\frac{1}{16} \cdot \frac{T_{i,cc} \cdot 10^3}{B_{cc}} \cdot 3n_i \cdot 2\pi a_{cc} L \right) \right. \quad (\text{A.50})$$

$$= \frac{8aB_{cc}}{3T_{i,cc}10^3} \quad (\text{A.51})$$

A.5.3 Gyro-Bohm diffusion

The gyro-Bohm scaling assumes cross-field transport is dominated by small ion-gyroscale turbulence. Though commonly used for tokamak scaling laws, we should be able to get some rough estimates for mirrors. Right now it can be estimated by just diving the Bohm confinement time by the normalized gyroradius ρ^* . The gyro-Bohm estimate is then:

$$\tau_{\text{gyro-Bohm}} = \frac{8aB_{cc}}{3T_{i,cc}10^3} \cdot \frac{1}{\rho^*} \quad (\text{A.52})$$

The $\frac{1}{\rho^*}$ factor can boost the confinement time estimate by a factor of 50-100. phil: But the confinement times still seems a little low.

A.5.4 ETG-driven transport

Cary's spreadsheet says phil: (I'm having trouble finding a derivation or hand-wavy justification)

:

$$\chi_{\text{ETG}} = 0.1 \frac{T_{e,cc}^{3/2}}{B_{cc}} \quad (\text{A.53})$$

$$\tau_{\text{ETG}} = \frac{a_{cc}^2}{\chi_{\text{ETG}}} \quad (\text{A.54})$$

A.6 Mirror-specific derived quantities

A.6.1 Temperatures and confinement time in a beam-heated mirror from Egedal et al 2022 [EEF22]

Electron and ion temperature (keV) via pure beam heating [EEF22]: we must solve a system of equations which considers the power balance of the machine. The ion temperature, given by eq. 22 in [EEF22] is:

$$\frac{3}{2} \frac{T_i}{E_{beam}} = \frac{\exp(-\alpha) - \alpha \Gamma(0, \alpha)}{\Gamma(0, \alpha)} \quad (\text{A.55})$$

Note that eq. 22 in [EEF22] is missing a factor of α in the numerator in front of the Γ function. The electron temperature can be found in terms of T_i and α by rearranging the definition of α (eq 21):

$$\frac{T_e}{E_{beam}} = \left(\frac{T_i}{E_{beam}} \frac{2}{3} \frac{\alpha^2 \ln R_m^2}{(22.4)^2} \right)^{1/3} \quad (\text{A.56})$$

These can be solved for with the help of a power (really energy-per-particle) balance equation (eq. 24 in [EEF22]):

$$E_{beam} + p_{aux} = T_i + 6T_e \quad (\text{A.57})$$

where p_{aux} is the combined sources and losses, such as alpha-particle heating, plasma heat losses, RF heating, radial transport, and so on. By balancing the power lost with auxiliary heating power we can keep $p_{aux} = 0$ to avoid iteratively solving this equation. Each ion loses $T_i + e\Phi \approx T_i + 5T_e$ units of energy and each electron loses $\approx T_e$ units because only hotter electrons can surmount the ambipolar potential. p_{aux} isn't so much a power as it is the energy gained/lost per particle – an actual power would require evaluation of the confinement time (eq 29 in [EEF22]):

$$\tau_p = \tau_{T_i}^{90} \frac{1}{\alpha_1 \lambda_1} \frac{\mathcal{H}}{T_i / E_{beam}} \int_0^1 M_1(\xi) d\xi \quad (\text{A.58})$$

where $\tau_{T_i}^{90}$ is the "scattering reactivity", α_1 , λ_1 , and M_1 are the normalization value, eigenvalue, and eigenfunction of the Lorentz scattering operator (eq 4 in [EEF22]). phil: I can't be bothered

to write a differentiable solver for the confinement time right now so we're just going to use the values of $\tau_{\text{Fowler-Baldwin}}$ for the first rough optimizations (which Cary says will give pessimistic estimates (actually, not sure)). By particle conservation and because NBI will be the dominant fueling mechanism, confinement time relates to density and beam current by:

$$\tau_p = eVn_b/I_{\text{NBI}} \quad (\text{A.59})$$

For ion temperature, Cary's spreadsheet says:

$$T_i = \frac{2}{3}E_{\text{inj}} \quad (\text{A.60})$$

This emerges from the relation that $E = \frac{1}{2}k_B T$ for every degree of freedom. For single particles, we assume 3 degrees of freedom to get $E = \frac{3}{2}T$ where T is expressed in eV. problem: If referencing Egedal et al 2022 [EEF22], this is only true if the electrons are very hot, i.e., $T_e(\log R_M)^{2/3}/E_{\text{inj}} \approx 1$. However, it is also shown in the paper that $T_i/E_{\text{inj}} \sim 0.6$. Assuming 2/3 is more conservative because a higher temperature means a lower reactivity and thus lower total fusion power at temperatures above 70-ish keV.

For electron temperature, Cary's spreadsheet says:

$$T_e = 0.089E_b \log_{10}(R_p)^{0.4} \quad (\text{A.61})$$

$$= 0.089E_b \log_{10}\left(\frac{R_p}{1-\beta}\right)^{0.4} \quad (\text{A.62})$$

which seems to give a roughly 2x higher electron temperature than the reduced model in Egedal 2022 [EEF22], which means that our estimate will be more optimistic. Kunal: Cary said this comes from working through the energy balance of a beam heated mirror device. Apparently it is in their new paper. I will go over this and figure it out. The $1 - \beta$ term is from the finite beta corrections to the mirror ratio.

Particle confinement time (Convention: $R_p = R_m$) found in Baldwin's end-loss paper [Bal77] equations 4.14 and 4.13. The same equation can be applied to tandem mirrors with thermal barriers

and plug cells [BL79]. According to Cary, this number will give pessimistic estimates. Equation 4.14 [Bal77] states:

$$n\tau_{\text{Fowler/Baldwin}} = \kappa \times 10^{10} E_b^{3/2} \log R_{\text{eff}} / \log 10 \quad (\text{A.63})$$

where $[n]$ is cm^{-3} and $[E_b]$ is keV, and

$$R_{\text{eff}} = R_m / (1 + (q\phi / mE_i)) \quad (\text{A.64})$$

For 90° NBI, κ falls between 2.4 and 2.8 according to Fokker-Planck calculations [Bal77]; it would be ~ 1.7 if the ion distribution did not have a loss-cone hole because the average energy is higher.

problem: Angled injection can impact this significantly (but ignoring for reduced model optimization) **phil:** I don't understand what this $(q\phi / mE_i)$ term is. This is a purely classical number – the main loss of ion energy is to electron drag, followed by ion-ion collisions / scattering into the loss cone (ions lost and accelerated by the ambipolar potential can be recaptured by direct energy conversion, but that is not accounted for here). Electrons are chilled by neutral beam injection and lost out the ends of the mirror if their energy exceeds the ambipolar potential.

Converting to $[n]$ in m^{-3} :

$$\tau_{\text{Fowler/Baldwin}} = 2.8 \cdot 10^{16} \frac{E_b^{3/2}}{n_e} \log R_m / \log 10 \quad (\text{A.65})$$

We may also need finite- β corrections to the mirror ratio.

A.6.2 Confinement time given by classical transport

Classical confinement time estimates assumes that transport is dominated by diffusion of gyrocenters via Coulomb collisions (from Chen section 5.8[Che16]). The diffusivity is:

$$D_{\text{classical}} = \eta_{\perp} n \sum T / B^2 \quad (\text{A.66})$$

where the perpindicular conductivity (for hydrogen) η_{\perp} is (temperatures in eV):

$$\eta_{\perp} = 2 \cdot \eta_{\parallel}, \quad (\text{A.67})$$

$$\eta_{\parallel} = 5.2 \cdot 10^{-5} \frac{Z \ln \Lambda_{ei}}{T_e^{3/2}} \sqrt{\mu} \quad (\text{A.68})$$

The confinement time is then (summing over species):

$$\tau_{\text{classical}} = \frac{nV}{A \cdot \Gamma} \quad (\text{A.69})$$

$$= \frac{na}{-2D_{\text{perp}} \nabla n} \quad (\text{A.70})$$

$$\tau_{\text{classical}} = \frac{aB^2}{-2\eta_{\perp} \nabla n \sum T} \quad (\text{A.71})$$

Again assuming a linear radial density profile with a peak of $3n_i$ to keep the total particle number $n_i \cdot V$:

$$\tau_{\text{classical}} = \frac{a^2 B^2 T_e^{\frac{3}{2}}}{3.12 \cdot 10^{-4} \cdot n Z \sqrt{\mu} \ln \Lambda \sum T} \quad (\text{A.72})$$

The aggregate confinement time is then:

$$\tau_{\text{tot}} = \frac{1}{\frac{1}{\tau_{\text{classical}}} + \frac{1}{\tau_{\text{Fowler/Bladwin}}} \quad (\text{A.73})}$$

A.6.3 End Cells/Plugs

Mirror ratio:

$$R_{\text{plug}} = \frac{B_{p,m}}{B_p} \quad (\text{A.74})$$

Radius at the midplane (mapped from bore radius):

$$a_{\text{plug}} = r_b \sqrt{\frac{B_{p,m}}{B_p}} \quad (\text{A.75})$$

Volume:

$$V_p = L_p \pi a_p^2 \quad (\text{A.76})$$

Total particle number:

$$N_{\text{tot}} = V_p n_{\text{plug}} \quad (\text{A.77})$$

Particles lost per second:

$$\frac{dN}{dt} = \frac{N_{\text{tot}}}{\tau_{\text{Fowler/Baldwin}}} \quad (\text{A.78})$$

Number of gyroradii in the plasma radius:

$$N_{\text{gyro}} = \frac{a_p}{\rho_i} \quad (\text{A.79})$$

Density (m^{-3}) at the β limit:

$$n_{\text{plug}} = B_p^2 \frac{\beta_{\text{limit}}}{2\mu_0|e|(T_{\text{ion}} + T_e)} \quad (\text{A.80})$$

Here, T_{ion} and T_e are expressed in eV. This can be found in Wesson page 115. Rolling all the constants together and with T_i and T_e in keV:

$$n_{20} = n_{\text{plug}}/10^{20} = B_p^2 \frac{\beta_{\text{limit}}}{0.04(T_{\text{ion}} + T_e)} \quad (\text{A.81})$$

NBI Current (A): phil: This should already account for dN/dt caused by fusion reactions if we assume that alphas have a similar confinement time? This is the current that the lost particles are reinjected after filtering out ash. Fusion reactions decrease the number of ions so N will actually be lower at a factor of around 1 – burnup fraction.

$$I_{\text{NBI}} = |e| \frac{dN}{dt} \quad (\text{A.82})$$

The neutral beam current is enough to replace the particles lost by the end plugs. In reality, this number will be larger since the beam neutrals are ionized via charge exchange as well as ion/electron impact.

Electron heating by fast ions (MW):

$$P_{\text{e heating by fast ions}} = 10^{-3} \frac{I_{\text{NBI}} E_b}{\tau_{\text{slow}}} \quad (\text{A.83})$$

Synchrotron radiation power loss (MW) [Wes87]:

$$P_{\text{synch}} = 6 \cdot 10^{-3} V_p n_{20} T_e \gamma^2 B_p^2 \quad (\text{A.84})$$

Bremsstrahlung radiation power loss (MW) [Wes87]:

$$P_{\text{brem}} = 5.35 \cdot 10^{-3} n_{20}^2 Z_{\text{eff}} \sqrt{T_e} V_p \quad (\text{A.85})$$

Power loss from escaping electrons (MW):

$$P_{e,\text{endloss}} = 10^{-3} (I_{\text{NBI}} + I_{\text{cooling}}) \cdot 7T_e \quad (\text{A.86})$$

I_{cooling} is non-zero when there is current in the expander/divertor. The $7T_e$ is because only electrons with an energy greater than the ambipolar potential can escape.

Power loss from escaping fast ions (MW):

$$P_{i,\text{endloss}} = 10^{-3} I_{\text{NBI}} (E_b - T_e) \quad (\text{A.87})$$

Injected NBI Power (MW):

$$P_{\text{NBI}} = 10^{-3} I_{\text{NBI}} E_b \quad (\text{A.88})$$

Injected ECH Power (MW): problem: Roll this into power balance equation (eq 24 in [EEF22]). But this will take considerable effort. phil: Electron endloss power should already be accounted for in the reduced model, but heating from fast ions is not. Synchotron + fast ion heating + Bremsstrahlung must be included in ECH to have a consistent T_e .

$$P_{\text{ECH}} = \frac{P_{\text{synch}}}{20} + P_{e,\text{endloss}} - (\text{Electron heating from fast ions}) \quad (\text{A.89})$$

Divide by 20 since the plasma recaptures most of the synchrotron losses are reabsorbed.

Lawson Triple Product ($10^{20}\text{keV}\cdot\text{s}/\text{m}^3$):

$$\tau_{\text{Fowler/Baldwin}} n_{20} T_i \quad (\text{A.90})$$

Neutron Flux (MW/m^2): problem: This should include DD and cat-DD neutrons.

$$\frac{14}{17.6} \frac{P_{\text{plug}}}{4\pi a_{\text{wall}}^2} \quad (\text{A.91})$$

phil: Does DD vs DT neutron flux significantly affect breeding ratios? For 400 keV ion temperatures it could lower the DT TBR requirements by like 10% or so – definitely significant when targeting TBRs of like 1.1. This info can be found on <https://www-nds.iaea.org/exfor/endf.htm>

using the targets LI-6; LI-7, reactions MT-105 and 205: N,T;N,XT, and extending the energy above 10 MeV.

phil: Burnup fraction, alpha particle density, and Zeff aren't really used in the optimization anywhere. It'd be tricky to include the effects of alpha particle density because that may change the slowing-down times which effects the power balance of the end plugs and so on. These effects would require time-evolution, which is beyond the scope of this "0D" analysis. We would expect the density and reaction rate error to be on the order of the burnup fraction, because ash can be filtered out and exhausted while the fuel is reinjected.

Burnup fraction:

$$\frac{R_{x,\text{plug,DT}}}{dN/dt} \quad (\text{A.92})$$

α particle density ($10^{20} m^{-3}$):

$$n_\alpha = \frac{I_{\text{NBI}} Q_{\text{plug}} \tau_\alpha E_b}{16 V_p E_\alpha} \quad (\text{A.93})$$

but a more intuitive way of putting it may be **phil:** (need to double check this) :

$$n_\alpha = \frac{\tau_\alpha (R_{x, \text{DT}} + \frac{1}{2} R_{x, \text{DD}})}{V} \quad (\text{A.94})$$

Z_{eff} : (from Wesson section 2.16 [Wes87]) **phil:** We assume no impurities! :

$$Z_{\text{eff}} = \frac{\sum_j n_j Z_j^2}{\sum_j n_j Z_j} = \frac{n + 4n_\alpha}{n + 2n_\alpha} \quad (\text{A.95})$$

Q_{plug} :

$$Q_{\text{plug}} = \frac{P_{\text{plug}}}{P_{\text{injected}}} \quad (\text{A.96})$$

phil: These quantities below are for a simple mirror. These will be duplicated for a tandem system and will be unused in any optimization since simple mirrors are unlikely to make a compelling reactor.

$P_{\text{electric,in}}$:

$$P_{\text{electric,in}} = P_{\text{total}} \left(\frac{1}{\eta_{HS}} - \eta_{DC} \left(1 - \frac{T_e}{E_b} \right) \right) \quad (\text{A.97})$$

$P_{\text{electric,out}}$: phil: The 0.8 represents the heating contribution from neutrons.

$$P_{\text{electric,out}} = 0.8 \eta_{HS} P_{\text{plug}} \quad (\text{A.98})$$

Q^* : phil: The 0.2 represents the alpha particle contribution. Note that

$$Q^* = \frac{Q_{\text{plug}}}{\frac{1}{\eta_{HS}} - \eta_{DC} \left(1 - \frac{T_e}{E_b} + 0.2 Q_{\text{plug}} \right)} \quad (\text{A.99})$$

Q_{electric} :

$$Q_{\text{electric}} = Q^* \cdot 0.8 \cdot \eta_{HS} \quad (\text{A.100})$$

A.6.4 Tandem mirror central cell

Radius at the midplane:

$$a_{cc} = r_b \sqrt{\frac{B_{p,m}}{B_{cc}}} \quad (\text{A.101})$$

Central cell mirror ratio:

$$R_{cc} = \frac{B_{p,m}}{B_{cc}} \quad (\text{A.102})$$

Central cell beta:

$$\beta_{cc} = \frac{2\mu_0|e|n_{cc}(T_{cc,i} + T_{cc,e})}{B_{cc}^2} \quad (\text{A.103})$$

$\beta_{cc} \geq 1$ will lead to an infinite Pastukhov factor, so the β -enhanced mirror ratio $R_{cc,\text{eff}} = R_{cc} \left(\sqrt{1 - \beta_{cc}} \right)^{-\frac{1}{2}}$

will be limited by keeping $\beta_{cc} \leq 0.9$.

In a tandem mirror (without a thermal barrier), we assume that the central cell electrons and plug cell electrons are Maxwellian and in thermal equilibrium, and that the central cell ions are also at the same temperature (Introduction to Tandem Mirror Physics, eq 1-3 (pg 78)):

$$T_{cc,i} = T_{cc,e} = T_{\text{plug},e} \cdot T_{\text{fudge factor}} \quad (\text{A.104})$$

The plug cell electron temperature is reduced by some fudge factor because they are heating the central cell plasma. Since the electrons follow a Maxwellian distribution along field lines, they

follow the Maxwell-Boltzmann relationship, where the potential difference between the plug and central cells are given by:

$$\Phi_i = \Phi_p - \Phi_c = T_{ep} \ln \left(\frac{n_p}{n_c} \right) \quad (\text{A.105})$$

The enhancement in ion confinement time in the central cell is then given by the Pastukhov factor (Pastukhov 1974, eq. 21 [Pas74], Kesner et al. eqs. 1-3 [KGL83]):

$$n_c \tau_i = n_c \tau_{ii} g(R) \frac{\Phi_i}{T_{ic}} \exp \left(\frac{\Phi_i}{T_{ic}} \right) \quad (\text{A.106})$$

where $g(R)$ is a weak function of the mirror ratio. We assume the $g(R)$ is:

$$g(R) = \log \left(2R_{cc} \frac{1}{\sqrt{1-\beta_{cc}}} + 1 \right) \quad (\text{A.107})$$

The ion confinement time is then:

$$\tau_E = \text{Pastukhov} \cdot \tau_{cc,ii} \quad (\text{A.108})$$

$$= \log \left(2R_{cc} \frac{1}{\sqrt{1-\beta_{cc}}} + 1 \right) \frac{T_{ep}}{T_{ic}} \ln \left(\frac{n_p}{n_{cc}} \right) \left(\frac{n_p}{n_{cc}} \right)^{T_{p,e}/T_{c,i}} \cdot \tau_{cc,ii} \quad (\text{A.109})$$

Since $T_{p,e} = T_{c,i}$, this reduces to

$$\tau_E = \log \left(2R_{cc} \frac{1}{\sqrt{1-\beta_{cc}}} + 1 \right) \ln \left(\frac{n_p}{n_{cc}} \right) \left(\frac{n_p}{n_{cc}} \right) \cdot \tau_{cc,ii} \quad (\text{A.110})$$

Thermal barriers are not considered in this analysis, which enhance the central cell confinement by elevating plug electron temperatures instead of only modifying the plug-central cell density ratio (see Post 1987 eq. 10-110[Pos87]). Thermal barriers require additional heating and ion pumpout methods. If estimates of the power requirements of thermal barriers are available, they can be easily included in this analysis and optimization process.

Power lost from the reactor by central cell particles, per meter (MW, T in keV):

$$P_{cc,loss} = 10^{-3} \pi \cdot a_{cc}^2 n_{cc} \cdot e \frac{3}{2} (T_{cc,i} + T_{cc,e}) / \tau_E \quad (\text{A.111})$$

Since this is axial power lost, it's assumed that this power (at least the ion contribution) is recovered by the DECs.

The power lost can be account for by lowering T_e by some fudge factor, or re-heating the electrons back up to the self-consistent temperature by injecting ECH:

$$P_{aux,ECH} = P_{cc,loss} \quad (\text{A.112})$$

The central cell will be fueld using cold gas puffing and is ionized and heated by electrons from the plugs. The fueling current is then:

$$I_{cc,fuel} = \frac{dN_{cc}}{dt} = \pi a_{cc}^2 L_{cc} n_{cc} / \tau_E \quad (\text{A.113})$$

Fusion Power per meter (MW/m): phil: here for legacy reasons. In the actual optimization procedure length will be one of the quantities that is optimized

$$P_{fusion} = 17.6 |e| R_x \quad (\text{A.114})$$

Breakeven length:

$$L_{breakeven} = \frac{2P_{\text{plug,injected}}}{P_{\text{fusion per m}}} \quad (\text{A.115})$$

phil: Cary's spreadsheet solves for L_{cc} given Q but we probably won't want Q directly in the cost function since we'll be optimizing for dollar cost or something else that depends on Q .

Central cell length:

$$L_{cc} = Q \cdot L_{breakeven} \quad (\text{A.116})$$

Total fusion power (MW):

$$P_{\text{total}} = 2P_{\text{plug}} + L_{cc} P_{\text{fusion}} \quad (\text{A.117})$$

A.6.5 Overall power balance and plant power estimates

Total electric power in:

$$P_{\text{electric,in}} = \eta_{ECH} P_{ECH} + \eta_{NBI} P_{NBI} + \eta_{RF} P_{RF} \quad (\text{A.118})$$

Recirculating power:

$$P_{\text{recirculating}} = \eta_{DEC} (P_{\text{fusion,charged}} + P_{cc,i,\text{endloss}} + P_{\text{plug,i,endloss}}) \quad (\text{A.119})$$

Thermal power, ignoring power generated by the blanket (the last term is thermal losses caused by DEC inefficiencies):

$$P_{\text{thermal}} = P_{\text{fusion,neutrons}} + (1 - \eta_{DEC}) \left(\frac{P_{\text{recirculating}}}{\eta_{DEC}} \right) \quad (\text{A.120})$$

Net electric power:

$$P_{\text{electric,net}} = -P_{\text{electric,in}} + P_{\text{recirculating}} + \eta_{\text{thermal}} P_{\text{thermal}} \quad (\text{A.121})$$

Q electric:

$$Q_{\text{electric}} = \frac{P_{\text{recirculating}} + \eta_{\text{thermal}} P_{\text{thermal}}}{P_{\text{electric,in}}} \quad (\text{A.122})$$

A.6.6 Instabilities

DCLC ratio (need to keep $\sim 1,000$) [KCP17, Pos66]:

$$\text{DCLC ratio} = \left(\frac{f_{pi}}{f_{ci,D}} \right)^2 \quad (\text{A.123})$$

The DCLC ratio must be kept $\sim 1,000$ as the radial density gradient needed to trigger the DCLC instability is very small ($I_{\text{gradient}} < 0.01\rho_{g,i}$ for stability). The above condition keeps the plasma radius large enough to prevent radial gradients that are sharper than those needed to trigger the DCLC instability from forming.

Interchange growth rate (s^{-1}):

$$\gamma_{\text{interchange}} = \frac{v_{Ti}}{L_p} \quad (\text{A.124})$$

Electron temperature gradient... something problem: need to understand

$$\chi_{\text{ETG}} = 0.1 \frac{T_{cc,e}^{3/2}}{B_{cc}} \quad (\text{A.125})$$

$$\tau_{\text{ETG}} = \frac{a_{cc}^2}{\chi_{\text{ETG}}} \quad (\text{A.126})$$

A.7 Costs and economics

A.7.1 Heating

problem: **Citations needed**

ECH: \$10/W

RF: \$1/W

NBI: \$5/W

A.7.2 Magnets

problem: **Citations needed**

kA-turns of coil needed for a given field and radius:

$$I_{\text{kA-turns}} = \frac{2 \cdot B \cdot a}{1000 \cdot \mu_0} \quad (\text{A.127})$$

kA-m of superconductor needed:

$$S = 2\pi R \cdot I_{\text{kA-turns}} \quad (\text{A.128})$$

Cost per kA·m = 10^{-4} M\$ / kA·m

Cost of magnet = $S \cdot (\text{cost per kA}\cdot\text{m})$

Radii of magnet coils needed: phil: The numbers below are from Cary's spreadsheet. I can't quite follow the thought process that went into these – the ones with my highlight are the quantities I'm using. problem: Rethink and justify these.

1. Mirror: $r_{\text{bore}} + d_{\text{vv}}$ (0.1m) + d_{blanket} (0.6m)

phil: $r_{\text{bore}} \cdot a_{\text{wall, ratio}} + d_{\text{vv}}$ (0.1m) + d_{blanket} (0.6m)

2. Plug midplane: $(a_{\text{wall, ratio}} \cdot a_{\text{plasma}}) + d_{\text{blanket}} + d_{\text{vv}}$ (0.2m)

phil: This will need to change depending on the length of the plug: we may need a solenoid or Maxwell coil instead of just a simple coil to keep

3. Plug divertor: beta limit + 0.2 problem: Doesn't make sense – ignoring for now

4. Central cell: $(a_{\text{wall, ratio}} \cdot a_{\text{plasma}}) + d_{\text{blanket}}$

phil: $(a_{\text{wall, ratio}} \cdot a_{\text{plasma}}) + d_{\text{vv}} + d_{\text{blanket}}$

For the central cell solenoid we are assuming a spacing of one coil per meter for diagnostic access. This is an adjustable parameter but will not be optimized because that would require energetic particle confinement estimates for coil ripple. For reference, the MARS study [LPG] had 42 central cell magnets spaced 3.16m apart with an inner radius of roughly 2m which led to 6% field ripple, which I assume is tolerable.

1. Solenoid field: $B = \mu_0 \cdot n_{\text{cc,turns}} \cdot I$, where $n_{\text{cc,turns}}$ is number of turns per coil. This becomes

$$B = \mu_0 \cdot I_{\text{kA-turns}}$$

2. kA-m per meter length (or per coil): $S_{\text{cc}} = 2\pi a_{\text{cc}} \cdot \frac{B}{\mu_0} \cdot (1/\text{coil spacing})$

A.8 Optimization constraints

phil: This method doesn't work for constraining the minimum field!

A.8.1 Midplane fields regularization via alpha particle confinement penalties

If we do not regularize field strengths, then the optimizer will bring the central cell (or plug) magnetic fields to 0 or negative. Only the midplane fields of the central cell and plugs will be regularized because the cost functions of interest tend towards higher reactor performance (and/or lower cost), and thus higher mirror ratios (and less HTS tape). The vacuum vessel should be, at minimum, four alpha gyroradii across. If an alpha is produced in the core, it will reach a distance

of two gyroradii if all the energy is perpendicular to the field (aside: this is more likely with spin-polarized fuels). Doubling the vacuum vessel radius to four alpha gyroradii is the safer bet. The 3.5 MeV alpha gyroradius is:

$$r_{\text{Larmor}} = \frac{\sqrt{2mE_{\perp}}}{2eB} = \frac{0.2694\text{cm}}{B} \quad (\text{A.129})$$

This regularization is enforced as a penalty coefficient on charged particle fusion power as an exponential function of the vessel wall:

$$\mathcal{C}_{\text{power penalty}} = \begin{cases} e^{a_{\text{diff}}/r_{\text{Larmor}}} & \text{if } a_{\text{diff}} > 0 \\ 1.0 & \text{otherwise} \end{cases} \quad (\text{A.130})$$

where a_{diff} is the difference between the vessel wall and 4 alpha gyroradii: $a_{\text{diff}} = 4r_{\text{Larmor}}$, 3.5 MeV alphas – a_{vv} . These particle losses depend on the radial plasma profile and should be simulated and implicitly affect the optimization instead of the explicit penalty as done here. Only the 3.5 MeV alpha gyroradius is considered because it's the largest of all the usual fusion products but we apply the penalty to all fusion products. This penalty aims to be a conservative estimate.

A.8.2 Kunal's suggestions

Kunal: If we want to pursue Cary's NBI only reactor design concept, then we need to use $\geq 100\text{keV}$ beams since the DT reaction cross section peaks at a center of mass energy of $\sim 65\text{keV}$. Currently, Cary is using 1MeV beams in his code.

I think we should just assume we are shooting for something that is Q=10 and has a usable power out of 200MW. This should stop our coding analysis from making something dumb like a Q=1,000 reactor which has a total power out of 5W by just using a very low beam current.

Cary calculates a lot of stuff with regards to the growth rate of various instabilities. Do we have/want to take all/some of the them into account in our optimization? For example, keeping the DCLC parameter at 1,000,

Per Kunal's suggestion we'll be operating under the following constraints:

1. Minimum usable power out- 200MW
2. Maximum NBI energy- 1MeV
3. Maximum Central cell length- 300m
4. Maximum plasma radius- 0.6m
5. Maximum field strength- 25T
6. Maximum beta- 0.8
7. Minimum DCLC ratio- 1000

These constraints will be soft either a quadratic or exponential penalty for exceeding them so that the cost function is differentiable.

A.9 Optimizing mirror configurations

A.9.1 Gradient descent using SymPy and JAX

Optimization is performed via gradient descent, that is, taking the gradient of some cost function \mathcal{C} with respect to some input parameter vector \vec{x} :

$$\vec{x} := \vec{x} - \nabla_{\vec{x}} \mathcal{C} \cdot \lambda \quad (\text{A.131})$$

where λ is the step size. Specific input values can be frozen by multiplying the gradient by a mask.

Equations are defined in SymPy, which are then lambdified to JAX expressions and then compiled by JAX's just-in-time (JIT) compiler on first run, or when `jax.jit` is called. JAX [jax] calculates the gradients of \mathcal{C} with respect to \vec{x} automatically. The step size λ may be tuned; larger step sizes may not be able to be used because propagating gradients through exponential functions in the temperature calculations can be unstable. We also use 64-bit floats so that large values of α (in the reduced temperature model from Egedal 2022 [EEF22]) remain calculable.

A.9.2 Example: optimizing Q in a simple mirror

As an example of a simple optimization task, we optimize to increase the Q of a simple mirror with classical radial transport. In this case, Q is just fusion power over NBI and ECH power. ECH power is only used to replace Bremsstrahlung and electron cyclotron losses to maintain self-consistent temperatures without requiring iterative solving. D-D fusion products are assumed to be burned instantly, though this only increases fusion power by roughly 7%.

Because the optimal solution is to decrease B_p until the mirror ratio explodes, we will add a $1/B_p$ penalty term to keep values reasonable. The cost function is then:

$$\mathcal{C} = -Q + 1/B_p \quad (\text{A.132})$$

This cost function has no meaningful physical interpretation.

For this optimization case, we hold constant auxiliary heating power ($p_{aux} = 0$ MW), plasma beta ($\beta = 0.8$), mirror bore radius ($r_b = 0.25$ m), length ($L_p = 20$ m), tritium fraction ($T_{\text{frac}} = 0.5$), and beam energy ($E_b = 1000$ keV) and optimize only the mirror field (B_{pm} , T) and central (midplane) field (B_p , T). Z_{eff} is assumed to be 1. B_p is initialized to 6 T, and eight different values of B_{pm} are initialized between 7 and 20 T.

In this optimization, the step size λ is set to 1. The optimization was run for 1000 steps which was chosen arbitrarily it doesn't converge in that step range (and we don't expect it to in this case).

Plots of the cost function \mathcal{C} and the gradient L2 norm for each different configuration can be seen in fig. A.4. The effects of the optimization on the fields B_{pm} and B_p can be seen in fig. A.5. The optimization favors lowering B_p until the regularization cost becomes significant at around step 60. The dramatic increase in mirror ratio leads to greater axial confinement, which decreases NBI current and power, leading to increased Q and decreased fusion power. Plots of Q and fusion power can be seen in fig. A.6. The effects of this optimization on the temperatures (or average energy in the ion case) can be seen in fig. A.7. The increased confinement time allows the beam ions more time to slow on the background electrons, decreasing T_i and increase T_e . The decreased T_i decreases D-D reactivity but *increases* D-T reactivity at a faster rate, leading to

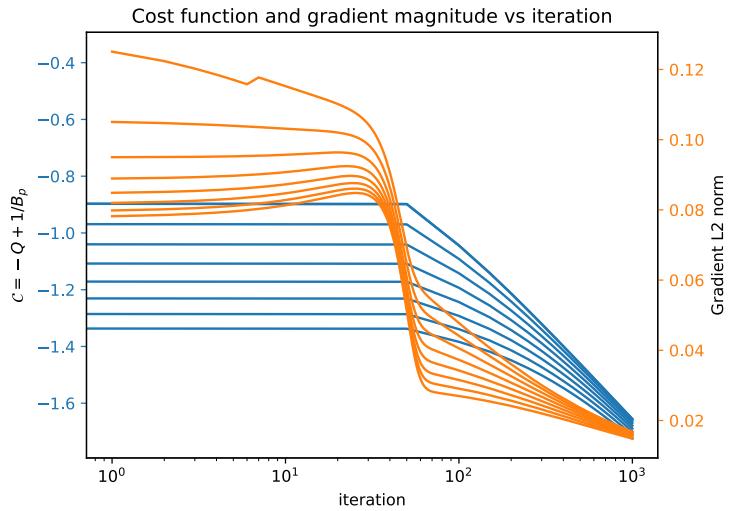


Figure A.4: The cost function and gradient magnitude for each optimization step.

higher fusion power. However, the lower density caused by the lower midplane field (as mandated by the β limit) causes a net *decrease* in fusion power. phil: The increased axial confinement time implies increased confinement of fusion alphas. The estimated total fuel burnup fraction $f_{\text{burnup}} = \frac{2R_x(\text{DT+DD})}{\frac{dN}{dr} + 2R_x(\text{DT+DD})}$ is 24% which may lower fusion power by roughly 42% (since the density would be 24% lower?). Something to think about.

Two insights can be gleaned from this simplified optimization task. Firstly, given optimistic physics, excessively high beam energies, incredibly high field strengths, and ignored impurity and ash accumulation, the reactor still only tops out at Q of around 2.3. This low Q implies that simple mirrors will never be a viable source of electricity. Secondly, Q is a shockingly bad optimization target because it maximizes fusion power *and* minimizes heating power simultaneously, thus high Q's can be obtained at low fusion power as demonstrated here. An expensive, low-power reactor is not useful.

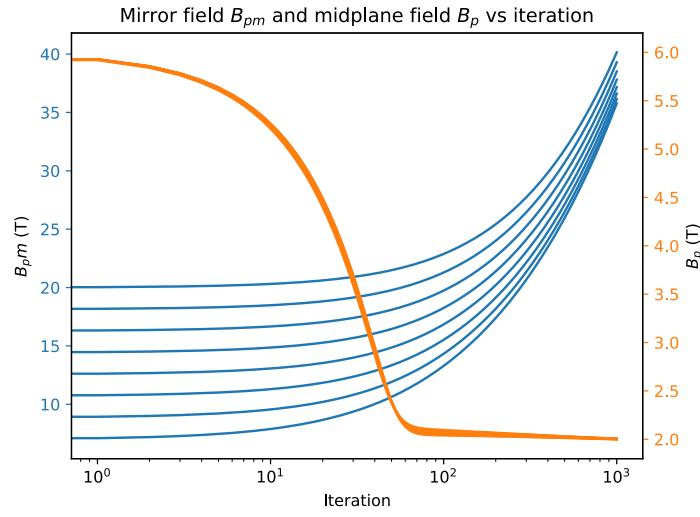


Figure A.5: The mirror and midplane fields for each optimization step. Note the logarithmic x-axis the rate of increase of the mirror field B_{pm} with respect to optimization step decreases with iteration.

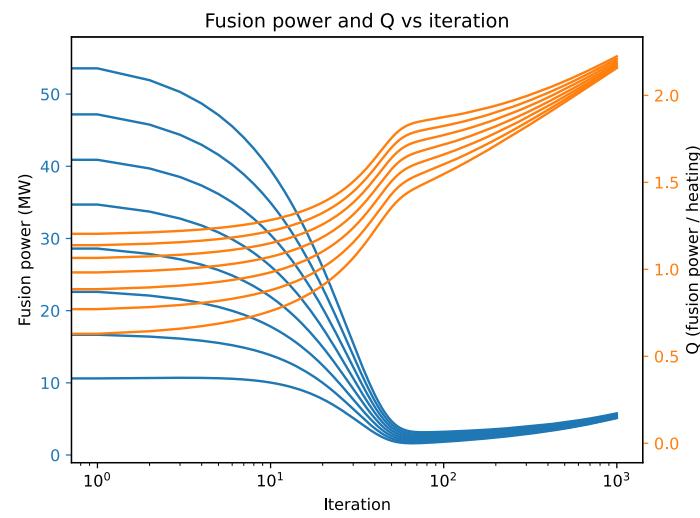


Figure A.6: The total fusion power and Q for each optimization step.

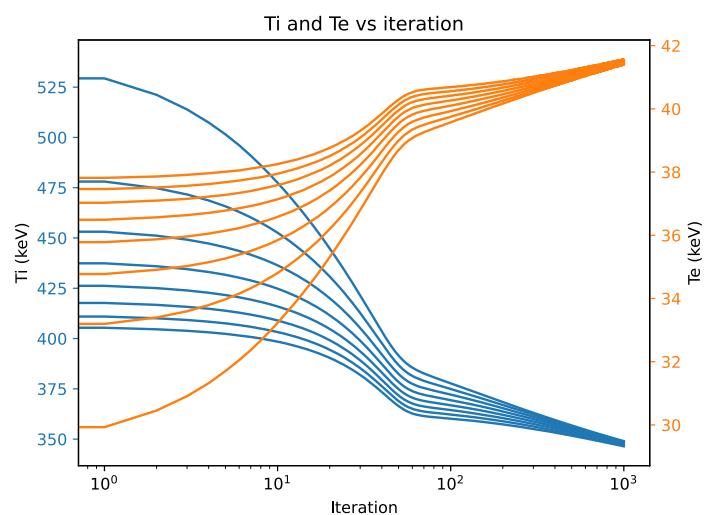


Figure A.7: The ion and electron temperatures for each optimization step

APPENDIX B

Automated Langmuir sweep analysis using machine learning

APPENDIX C

Wisdom acquired

1. Simple stuff goes a long way. Try it first
2. Listen to your advisor
3. Write up as you go along. It helps with the thinking too
4. Wrap up projects (or at least get them to a terminal state) as soon as possible
5. Get a dog
6. Nothing really matters
7. Have two simultaneous projects ongoing, not more, not less
8. Good feedback is hard to find. Seek it out
9. Take risks (see item 6)
10. Don't do solo projects
11. Make lots of friends
12. When life gives you lemons, give them to Mel to make lemon bars

REFERENCES

- [AAA23] Rushil Anirudh, Rick Archibald, M. Salman Asif, Markus M. Becker, Sadruddin Benkadda, Peer-Timo Bremer, Rick H. S. Budé, C. S. Chang, Lei Chen, R. M. Churchill, Jonathan Citrin, Jim A. Gaffney, Ana Gainaru, Walter Gekelman, Tom Gibbs, Satoshi Hamaguchi, Christian Hill, Kelli Humbird, Sören Jalas, Satoru Kawaguchi, Gon-Ho Kim, Manuel Kirchen, Scott Klasky, John L. Kline, Karl Krushelnick, Bogdan Kustowski, Giovanni Lapenta, Wenting Li, Tammy Ma, Nigel J. Mason, Ali Mesbah, Craig Michoski, Todd Munson, Izumi Murakami, Habib N. Najm, K. Erik J. Olofsson, Seolhye Park, J. Luc Peterson, Michael Probst, David Pugmire, Brian Sammuli, Kapil Sawlani, Alexander Scheinker, David P. Schissel, Rob J. Shalloo, Jun Shinagawa, Jaegu Seong, Brian K. Spears, Jonathan Tennyson, Jayaraman Thiagarajan, Catalin M. Tico, Jan Trieschmann, Jan Van Dijk, Brian Van Essen, Peter Ventzek, Haimin Wang, Jason T. L. Wang, Zhehui Wang, Kristian Wende, Xueqiao Xu, Hiroshi Yamada, Tatsuya Yokoyama, and Xinhua Zhang. “2022 Review of Data-Driven Plasma Science.” *IEEE Transactions on Plasma Science*, **51**(7):1750–1838, July 2023.
- [ACK21] J. Abbate, R. Conlin, and E. Kolemen. “Data-driven profile prediction for DIII-D.” *Nuclear Fusion*, **61**(4):046027, April 2021.
- [Bal77] D. E. Baldwin. “End-loss processes from mirror machines.” *Reviews of Modern Physics*, **49**(2):317–339, April 1977.
- [BBC10] Alexei D. Beklemishev, Peter A. Bagryansky, Maxim S. Chaschin, and Elena I. Soldatkina. “Vortex Confinement of Plasmas in Symmetric Mirror Traps.” *Fusion Science and Technology*, **57**(4):351–360, May 2010.
- [BBS07] P.A. Bagryansky, A.D. Beklemishev, and E.I. Soldatkina. “Influence of Radial Electric Field on High-Beta Plasma Confinement in the Gas Dynamic Trap.” *Fusion Science and Technology*, **51**(2T):340–342, February 2007.
- [BGB05] F. Brochard, E. Gravier, and G. Bonhomme. “Transition from flute modes to drift waves in a magnetized plasma column.” *Physics of Plasmas*, **12**(6):062104, June 2005.
- [BH92] H.-S Bosch and G.M Hale. “Improved formulas for fusion cross-sections and thermal reactivities.” *Nuclear Fusion*, **32**(4):611–631, 4 1992.
- [BKP82] J. M. Beall, Y. C. Kim, and E. J. Powers. “Estimation of wavenumber and frequency spectra using fixed probe pairs.” *Journal of Applied Physics*, **53**(6):3933–3940, 06 1982.
- [BL79] D. E. Baldwin and B. G. Logan. “Improved Tandem Mirror Fusion Reactor.” *Phys. Rev. Lett.*, **43**:1318–1321, 10 1979.

- [BLZ03] P.A. Bagryansky, A.A. Lizunov, A.A. Zuev, E. Yu. Kolesnikov, and A.L. Solomachin. “Experiments with Controllable Application of Radial Electric Fields in GDT Central Cell.” *Fusion Science and Technology*, **43**(1T):152–156, January 2003.
- [BRT91] H. L. Berk, D. D. Ryutov, and Yu. A. Tsidulko. “Temperature-gradient instability induced by conducting end walls.” *Physics of Fluids B: Plasma Physics*, **3**(6):1346–1354, June 1991.
- [CCA24] Ian Char, Youngseog Chung, Joseph Abbate, Egemen Kolemen, and Jeff Schneider. “Full Shot Predictions for the DIII-D Tokamak via Deep Recurrent Networks.”, April 2024. arXiv:2404.12416 [physics].
- [Che16] F. Chen. *Introduction to Plasma Physics and Controlled Fusion*. Springer, third edition, 2016.
- [CM09] T. A. Carter and J. E. Maggs. “Modifications of turbulence and turbulent transport associated with a bias-induced confinement transition in the Large Plasma Device.” *Physics of Plasmas*, **16**(1):012304, January 2009.
- [CS82] T. A. Casper and Gary R. Smith. “Observation of Alfvén Ion-Cyclotron Fluctuations in the End-Cell Plasma in the Tandem Mirror Experiment.” *Phys. Rev. Lett.*, **48**:1015–1018, Apr 1982.
- [CYK05] T. Cho, M. Yoshida, J. Kohagura, M. Hirata, T. Numakura, H. Higaki, H. Hojo, M. Ichimura, K. Ishii, K. Md. Islam, A. Itakura, I. Katanuma, Y. Nakashima, T. Saito, Y. Tatematsu, M. Yoshikawa, Y. Kojima, S. Tokioka, N. Yokoyama, Y. Tomii, T. Imai, V. P. Pastukhov, and S. Miyoshi. “Observation of the Effects of Radially Sheared Electric Fields on the Suppression of Turbulent Vortex Structures and the Associated Transverse Loss in GAMMA 10.” *Physical Review Letters*, **94**(8):085002, March 2005.
- [DEH23] Andreas Döpp, Christoph Eberle, Sunny Howard, Faran Irshad, Jinpu Lin, and Matthew Streeter. “Data-driven science and machine learning methods in laserplasma physics.” *High Power Laser Science and Engineering*, **11**:e55, 2023.
- [DFH23] G A Daly, J E Fieldsend, G Hassall, and G R Tabor. “Data-driven plasma modelling: surrogate collisional radiative models of fluorocarbon plasmas from deep generative autoencoders.” *Machine Learning: Science and Technology*, **4**(3):035035, September 2023.
- [DLD21] Jiaolong Dong, Jianchao Li, Yonghua Ding, Xiaoqing Zhang, Nengchao Wang, Da Li, Wei Yan, Chengshuo Shen, Ying He, Xiehang Ren, and Donghui Xia. “Machine learning application to predict the electron temperature on the J-TEXT tokamak.” *Plasma Science and Technology*, **23**(8):085101, August 2021.
- [Dol82] Thomas James Dolan. *Fusion research*. Pergamon Press, 1982.

- [EAB23] D. Endrizzi, J.K. Anderson, M. Brown, J. Egedal, B. Geiger, R.W. Harvey, M. Ialovega, J. Kirch, E. Peterson, Yu.V. Petrov, J. Pizzo, T. Qian, K. Sanwalka, O. Schmitz, J. Wallace, D. Yakovlev, M. Yu, and C.B. Forest. “Physics basis for the Wisconsin HTS Axisymmetric Mirror (WHAM).” *Journal of Plasma Physics*, **89**(5):975890501, October 2023.
- [EEF22] J. Egedal, D. Endrizzi, C.B. Forest, and T.K. Fowler. “Fusion by beam ions in a low collisionality, high mirror ratio magnetic mirror.” *Nuclear Fusion*, **62**(12):126053, 11 2022.
- [EPC09] E. T. Everson, P. Pribyl, C. G. Constantin, A. Zylstra, D. Schaeffer, N. L. Kugland, and C. Niemann. “Design, construction, and calibration of a three-axis, high-frequency magnetic probe (B-dot probe) as a diagnostic for exploding plasmas.” *Review of Scientific Instruments*, **80**(11):113505, November 2009.
- [FAE24] C.B. Forest, J.K. Anderson, D. Endrizzi, J. Egedal, S. Frank, K. Furlong, M. Ialovega, J. Kirch, R.W. Harvey, B. Lindley, Yu.V. Petrov, J. Pizzo, T. Qian, K. Sanwalka, O. Schmitz, J. Wallace, D. Yakovlev, and M. Yu. “Prospects for a high-field, compact break-even axisymmetric mirror (BEAM) and applications.” *Journal of Plasma Physics*, **90**(1):975900101, February 2024.
- [FCU13] B. Friedman, T. A. Carter, M. V. Umansky, D. Schaffner, and I. Joseph. “Nonlinear instability in simulations of Large Plasma Device turbulence).” *Physics of Plasmas*, **20**(5):055704, 05 2013.
- [FEE20] Yichen Fu, David Eldon, Keith Erickson, Kornee Kleijwegt, Leonard Lupin-Jimenez, Mark D. Boyer, Nick Eidietis, Nathaniel Barbour, Olivier Izacard, and Egemen Kolenmen. “Machine learning control for disruption and tearing mode avoidance.” *Physics of Plasmas*, **27**(2):022501, February 2020.
- [FKR79] S. Fornaca, Y. Kiwamoto, and N. Rynn. “Experimental Stabilization of Interchange Mode by Surface Line Tying.” *Phys. Rev. Lett.*, **42**:772–776, Mar 1979.
- [FLM06] A. Fasoli, B. Labit, M. McGrath, S. H. Müller, G. Plyushchev, M. Podestà, and F. M. Poli. “Electrostatic turbulence and transport in a simple magnetized plasma.” *Physics of Plasmas*, **13**(5):055902, May 2006.
- [Fri13] Brett Cory Friedman. *Simulation Analysis of Zero Mean Flow Edge Turbulence in LAPD*. PhD thesis, 2013.
- [FWD83] John R. Ferron, Alfred Y. Wong, Guy Dimonte, and Bernard J. Leikind. “Interchange stability of an axisymmetric, average minimum- B magnetic mirror.” *The Physics of Fluids*, **26**(8):2227–2233, August 1983.
- [GH71] G. E. Guest and E. G. Harris. “Flute Stabilization Via Electrostatically Confined Cold Electrons.” *Phys. Rev. Lett.*, **27**:1500–1503, Nov 1971.

- [GKG22] S. Ghazaryan, M. Kaloyan, W. Gekelman, Z. Lucky, S. Vincena, S. K. P. Tripathi, P. Pribyl, and C. Niemann. “Thomson scattering on the large plasma device.” *Review of Scientific Instruments*, **93**(8):083514, August 2022.
- [Gol95] R.J. Goldston. *Introduction to Plasma Physics*. CRC Press, Boca Raton, 1995.
- [GPL16a] W. Gekelman, P. Pribyl, Z. Lucky, M. Drandell, D. Leneman, J. Maggs, S. Vincena, B. Van Compernolle, S. K. P. Tripathi, G. Morales, T. A. Carter, Y. Wang, and T. DeHaas. “The upgraded Large Plasma Device, a machine for studying frontier basic plasma physics.” *Review of Scientific Instruments*, **87**(2):025105, February 2016.
- [GPL16b] W. Gekelman, P. Pribyl, Z. Lucky, M. Drandell, D. Leneman, J. Maggs, S. Vincena, B. Van Compernolle, S. K. P. Tripathi, G. Morales, T. A. Carter, Y. Wang, and T. DeHaas. “The upgraded Large Plasma Device, a machine for studying frontier basic plasma physics.” *Review of Scientific Instruments*, **87**(2):025105, February 2016.
- [GPS17] Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. “On Calibration of Modern Neural Networks.” <http://arxiv.org/abs/1706.04599>, August 2017. arXiv:1706.04599 [cs].
- [HBF84] Jr. Hooper, E. B., D. E. Baldwin, T. K. Fowler, R. J. Kane, and W. C. Turner. “Radial transport reduction in tandem mirrors using endwall boundary conditions.” *The Physics of Fluids*, **27**(9):2264–2267, 09 1984.
- [Hor99] W. Horton. “Drift waves and transport.” *Reviews of Modern Physics*, **71**(3):735–778, April 1999.
- [HPC05] W. Horton, Jean C. Perez, Troy Carter, and Roger Bengtson. “Vorticity probes and the characterization of vortices in the KelvinHelmholtz instability in the large plasma device experiment.” *Physics of Plasmas*, **12**(2):022303, February 2005.
- [IP13] A A Ivanov and V V Prikhodko. “Gas-dynamic trap: an overview of the concept and experimental results.” *Plasma Physics and Controlled Fusion*, **55**(6):063001, may 2013.
- [JAC22] Azarakhsh Jalalvand, Joseph Abbate, Rory Conlin, Geert Verdoollaeghe, and Egemen Kolemen. “Real-Time and Adaptive Reservoir Computing With Application to Profile Prediction in Fusion Plasma.” *IEEE Transactions on Neural Networks and Learning Systems*, **33**(6):2630–2641, June 2022.
- [Jas72] D. L. Jassby. “Transverse Velocity Shear Instabilities within a Magnetically Confined Plasma.” *The Physics of Fluids*, **15**(9):1590–1604, September 1972.
- [jax] “JAX: High-Performance Array Computing.”

- [JKC24] Samuel Jackson, Saiful Khan, Nathan Cummings, James Hodson, Shaun De Witt, Stanislas Pamela, Rob Akers, and Jeyan Thiyyagalingam. “FAIR-MAST: A fusion device data management system.” *SoftwareX*, **27**:101869, September 2024.
- [Kan79] B.I. Kanaev. “Stabilization of drift loss-cone instability (DCI) by addition of cold ions.” *Nuclear Fusion*, **19**(3):347, mar 1979.
- [Kar19] Andrej Karpathy. “A Recipe for Training Neural Networks.” <https://web.archive.org/web/20240709000647/http://karpathy.github.io/2019/04/25/recipe/>, April 2019. Accessed: 2024-07-12.
- [KCP17] Igor A. Kotelnikov, Ivan S. Chernoshtanov, and Vadim V. Prikhodko. “Electrostatic instabilities in a mirror trap revisited.” *Physics of Plasmas*, **24**(12):122512, 12 2017.
- [KGL83] J Kesner, MJ Gerver, BG Lane, BD Vey, RE Aamodt, PJ Catto, DA Ippolito, and JR Myra. “Introduction to tandem mirror physics.” 1983.
- [KHS87] Akio Komori, Yoshiya Higuchi, Yusuke Suetsugu, Akira Yonesu, and Yoshinobu Kawai. “Flute Stabilization of a Mirror-Confined Plasma by a Positive Ambipolar Potential.” *Journal of the Physical Society of Japan*, **56**(8):2607–2610, 1987.
- [KLZ21] Igor Kotelnikov, Andrej Lizunov, and Qiusun Zeng. “On the stability of small-scale ballooning modes in axisymmetric mirror traps.” *Plasma Science and Technology*, **24**(1):015102, 11 2021.
- [KST19] Julian Kates-Harbeck, Alexey Svyatkovskiy, and William Tang. “Predicting disruptive instabilities in controlled fusion plasmas through deep learning.” *Nature*, **568**(7753):526–531, April 2019.
- [lhd] “LHD experiment data repository.” doi:10.57451/lhd.analyzed-data.
- [Lie85] Paulett C. Liewer. “Measurements of microturbulence in tokamaks and comparisons with theories of turbulence and anomalous transport.” *Nuclear Fusion*, **25**(5):543–621, May 1985.
- [LPB17] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. “Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles.” <http://arxiv.org/abs/1612.01474>, November 2017. arXiv:1612.01474 [cs, stat].
- [LPG] B G Logan, L J Perkins, and J D Gordon. “Mirror Advanced Reactor Study (MARS): executive summary and overview.”
- [MCT07] J. E. Maggs, T. A. Carter, and R. J. Taylor. “Transition from Bohm to classical diffusion due to edge rotation of a cylindrical plasma.” *Physics of Plasmas*, **14**(5):052507, 05 2007.

- [MII91] A. Mase, A. Itakura, M. Inutake, K. Ishii, J.H. Jeong, K. Hattori, and S. Miyoshi. “Control of the radial electric field and of turbulent fluctuations in a tandem mirror plasma.” *Nuclear Fusion*, **31**(9):1725, sep 1991.
- [MM97] J. E. Maggs and G. J. Morales. “Fluctuations associated with a filamentary density depletion.” *Physics of Plasmas*, **4**(2):290–299, February 1997.
- [MMC05] J. E. Maggs, G. J. Morales, and T. A. Carter. “An Alfvén wave maser in the laboratory.” *Physics of Plasmas*, **12**(1):013103, January 2005.
- [MPL20] Andrea Murari, Emmanuele Peluso, Michele Lungaroni, Riccardo Rossi, Michela Gel-fusa, and JET Contributors. “Investigating the Physics of Tokamak Global Stability with Interpretable Machine Learning Tools.” *Applied Sciences*, **10**(19):6683, September 2020.
- [MRP24] Andrew Maris, Cristina Rea, Alessandro Pau, Wenhui Hu, Bingjia Xiao, Robert Granetz, Earl Marmar, the EUROfusion Tokamak Exploitation team, the Alcator C-Mod team, the ASDEX Upgrade team, the DIII-D. team, the EAST team, and the TCV team. “Correlation of the L-mode density limit with edge collisionality.”, June 2024. arXiv:2406.18442 [physics].
- [NKB19] Preetum Nakkiran, Gal Kaplun, Yamini Bansal, Tristan Yang, Boaz Barak, and Ilya Sutskever. “Deep Double Descent: Where Bigger Models and More Data Hurt.” <http://arxiv.org/abs/1912.02292>, December 2019. arXiv:1912.02292 [cs, stat].
- [NW94] D.A. Nix and A.S. Weigend. “Estimating the mean and variance of the target probability distribution.” In *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN’94)*, pp. 55–60 vol.1, Orlando, FL, USA, 1994. IEEE.
- [Pas74] V.P. Pastukhov. “Collisional losses of electrons from an adiabatic trap in a plasma with a positive potential.” *Nuclear Fusion*, **14**(1):3–6, 1 1974.
- [PBD06] F. M. Poli, S. Brunner, A. Diallo, A. Fasoli, I. Furno, B. Labit, S. H. Müller, G. Plyushchev, and M. Podestà. “Experimental characterization of drift-interchange instabilities in a simple toroidal plasma.” *Physics of Plasmas*, **13**(10):102104, October 2006.
- [PMC22] Conor Perks, Saskia Mordijk, Troy Carter, Bart Van Compernolle, Stephen Vincena, Giovanni Rossi, and David Schaffner. “Impact of the electron density and temperature gradient on drift-wave turbulence in the Large Plasma Device.” *Journal of Plasma Physics*, **88**(4):905880405, August 2022.
- [PMK23] A Pavone, A Merlo, S Kwak, and J Svensson. “Machine learning and Bayesian inference in nuclear fusion research: an overview.” *Plasma Physics and Controlled Fusion*, **65**(5):053001, May 2023.

- [Pos66] R. F. Post. “Electrostatic instabilities in Finite Mirror-confined plasmas.” *Physics of Fluids*, **9**(4):730, 1966.
- [Pos87] R.F. Post. “The magnetic mirror approach to fusion.” *Nuclear Fusion*, **27**(10):1579, 10 1987.
- [Pow74] E.J. Powers. “Spectral techniques for experimental investigation of plasma diffusion due to polychromatic fluctuations.” *Nuclear Fusion*, **14**(5):749, nov 1974.
- [PUC10] P. Popovich, M. V. Umansky, T. A. Carter, and B. Friedman. “Analysis of plasma instabilities and verification of the BOUT code for the Large Plasma Device.” *Physics of Plasmas*, **17**(10):102107, October 2010.
- [QGP23a] Yuchen Qian, Walter Gekelman, Patrick Pribyl, Tom Sketchley, Shreekrishna Tripathi, Zoltan Lucky, Marvin Drandell, Stephen Vincena, Thomas Look, Phil Travis, Troy Carter, Gary Wan, Mattia Cattelan, Graeme Sabiston, Angelica Ottaviano, and Richard Wirz. “Design of the Lanthanum hexaboride based plasma source for the large plasma device at UCLA.” *Review of Scientific Instruments*, **94**(8):085104, 08 2023.
- [QGP23b] Yuchen Qian, Walter Gekelman, Patrick Pribyl, Tom Sketchley, Shreekrishna Tripathi, Zoltan Lucky, Marvin Drandell, Stephen Vincena, Thomas Look, Phil Travis, Troy Carter, Gary Wan, Mattia Cattelan, Graeme Sabiston, Angelica Ottaviano, and Richard Wirz. “Design of the Lanthanum hexaboride based plasma source for the large plasma device at UCLA.” *Review of Scientific Instruments*, **94**(8):085104, August 2023.
- [RBC11] D. D. Ryutov, H. L. Berk, B. I. Cohen, A. W. Molvik, and T. C. Simonen. “Magneto-hydrodynamically stable axisymmetric mirrors.” *Physics of Plasmas*, **18**(9):092301, 09 2011.
- [RME19] C. Rea, K.J. Montes, K.G. Erickson, R.S. Granetz, and R.A. Tinguely. “A real-time machine learning-based disruption predictor in DIII-D.” *Nuclear Fusion*, **59**(9):096016, September 2019.
- [RWB08] C.M. Roach, M. Walters, R.V. Budny, F. Imbeaux, T.W. Fredian, M. Greenwald, J.A. Stillerman, D.A. Alexander, J. Carlsson, J.R. Cary, F. Ryter, J. Stober, P. Goihil, C. Greenfield, M. Murakami, G. Bracco, B. Esposito, M. Romanelli, V. Parail, P. Stubberfield, I. Voitsekhovitch, C. Brickley, A.R. Field, Y. Sakamoto, T. Fujita, T. Fukuda, N. Hayashi, G.M.D Hogeweij, A. Chudnovskiy, N.A. Kinerva, C.E. Kessel, T. Aniel, G.T. Hoang, J. Ongena, E.J. Doyle, W.A. Houlberg, A.R. Polevoi, ITPA Confinement Database and Modelling Topical Group, and ITPA Transport Physics Topical Group. “The 2008 Public Release of the International Multi-tokamak Confinement Profile Database.” *Nuclear Fusion*, **48**(12):125001, December 2008.
- [Sch13] David Schaffner. *Study of Flow, Turbulence and Transport on the Large Plasma Device*. PhD thesis, 2013.

- [SCR12] D. A. Schaffner, T. A Carter, G. D. Rossi, D. S. Guice, J. E. Maggs, S. Vincena, and B. Friedman. “Modification of Turbulent Transport with Continuous Variation of Flow Shear in the Large Plasma Device.” *Physical Review Letters*, **109**(13):135002, September 2012.
- [SCR13] D. A. Schaffner, T. A. Carter, G. D. Rossi, D. S. Guice, J. E. Maggs, S. Vincena, and B. Friedman. “Turbulence and transport suppression scaling with flow shear on the Large Plasma Device.” *Physics of Plasmas*, **20**(5):055907, May 2013.
- [Sim76] T. C. Simonen. “Measurements of ion cyclotron instability characteristics in a mirror-confined plasma.” *The Physics of Fluids*, **19**(9):1365–1370, 09 1976.
- [SKJ24] Jaemin Seo, SangKyeun Kim, Azarakhsh Jalalvand, Rory Conlin, Andrew Rothstein, Joseph Abbate, Keith Erickson, Josiah Wai, Ricardo Shousha, and Egemen Kolemen. “Avoiding fusion plasma tearing instability with deep reinforcement learning.” *Nature*, **626**(8000):746–751, February 2024.
- [SKR23] Rylan Schaeffer, Mikail Khona, Zachary Robertson, Akhilan Boopathy, Kateryna Pitsunova, Jason W. Rocks, Ila Rani Fiete, and Oluwasanmi Koyejo. “Double Descent Demystified: Identifying, Interpreting & Ablating the Sources of a Deep Learning Puzzle.” <http://arxiv.org/abs/2303.14151>, March 2023. arXiv:2303.14151 [cs, stat].
- [SNK21] Jaemin Seo, Y.-S. Na, B. Kim, C.Y. Lee, M.S. Park, S.J. Park, and Y.H. Lee. “Feedforward beta control in the KSTAR tokamak by deep reinforcement learning.” *Nuclear Fusion*, **61**(10):106010, October 2021.
- [SNK22] J. Seo, Y.-S. Na, B. Kim, C.Y. Lee, M.S. Park, S.J. Park, and Y.H. Lee. “Development of an operation trajectory design algorithm for control of multiple 0D parameters using deep reinforcement learning in KSTAR.” *Nuclear Fusion*, **62**(8):086049, August 2022.
- [STA22] Maximilian Seitzer, Arash Tavakoli, Dimitrije Antic, and Georg Martius. “On the Pitfalls of Heteroscedastic Uncertainty Estimation with Probabilistic Neural Networks.” <http://arxiv.org/abs/2203.09168>, April 2022. arXiv:2203.09168 [cs, stat].
- [SWP20] Prem Seetharaman, Gordon Wichern, Bryan Pardo, and Jonathan Le Roux. “AutoClip: Adaptive Gradient Clipping for Source Separation Networks.” <http://arxiv.org/abs/2007.14469>, July 2020. arXiv:2007.14469 [cs, eess, stat].
- [TFM09] G R Tynan, A Fujisawa, and G McKee. “A review of experimental drift turbulence studies.” *Plasma Physics and Controlled Fusion*, **51**(11):113001, November 2009.
- [tra] “Turbulence and transport in mirror geometries in the Large Plasma Device.” **91**.
- [Tra25] Phil Travis. “physicistphil/lapd-isat-predict: 2025-3-11.”, March 2025. doi:10.5281/zenodo.15007853.

- [VG06] Stephen Vincena and Walter Gekelman. “Drift-Alfvén wave mediated particle transport in an elongated density depression.” *Physics of Plasmas*, **13**(6):064503, June 2006.
- [VGP11] B. Van Compernolle, W. Gekelman, P. Pribyl, and C. M. Cooper. “Wave and transport studies utilizing dense plasma filaments generated with a lanthanum hexaboride cathode.” *Physics of Plasmas*, **18**(12):123501, 12 2011.
- [VMD22] J. Vega, A. Murari, S. Dormido-Canto, G. A. Rattá, M. Gelfusa, and JET Contributors. “Disruption prediction with artificial intelligence techniques in tokamak plasmas.” *Nature Physics*, **18**(7):741–750, July 2022.
- [VS22] Matias Valdenegro-Toro and Daniel Saromo. “A Deeper Look into Aleatoric and Epistemic Uncertainty Disentanglement.” 2022. arXiv:2204.09308 [cs.LG].
- [Wes87] J. Wesson. *4.9 - Radiation losses*, p. 100101. Clarendon Press, 1987.
- [WH84] Masahiro Wakatani and Akira Hasegawa. “A collisional drift wave description of plasma edge turbulence.” *The Physics of Fluids*, **27**(3):611–618, 03 1984.
- [WV82] Michael Wickham and Guy Vandegrift. “Curvature-induced interchange mode in an axisymmetric plasma.” *The Physics of Fluids*, **25**(1):52–58, January 1982.
- [WYP22] Chenguang Wan, Zhi Yu, Alessandro Pau, Xiaojuan Liu, and Jiangang Li. “EAST discharge prediction without integrating simulation results.” *Nuclear Fusion*, **62**(12):126060, December 2022.
- [YMM10] M. Yoshikawa, Y. Miyata, M. Mizuguchi, N. Imai, H. Hojo, M. Ichimura, T. Kariya, I. Katanuma, Y. Nakashima, R. Minami, H. Shidara, Y. Yamaguchi, Y. Shima, Y. Ohno, F. Yaguchi, and T. Imai. “Use of a Gold Neutral Beam Probe to Study Fluctuation Suppression During Potential Formation in the GAMMA 10 Tandem Mirror.” *Fusion Science and Technology*, **57**(4):312–319, May 2010.