

Development of a Time-Resolved Laser-Induced Fluorescence Technique for Nonperiodic Oscillations

by

Christopher J. Durot

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied Physics)
in the University of Michigan
2016

Doctoral Committee:

Professor Alec D. Gallimore, Chair

Professor John E. Foster

Assistant Research Scientist Timothy B. Smith

Professor Azer P. Yalin, Colorado State University

© Christopher J. Durot 2016

All Rights Reserved

To my grandmothers, Rose and Emma, whose early influence helped put me on this path.

Acknowledgements

First and foremost, I gratefully acknowledge my adviser, Alec Gallimore. He built the lab up from a shell, albeit a $6\text{ m} \times 9\text{ m}$ shell, into one of the foremost electric propulsion laboratories in the world. Thank you for providing me and all of your students with opportunities to conduct world-class research simply not available anywhere else, as well as for generously providing your knowledge and experience whenever it is needed. You balance lightly pushing to get things done while allowing the time and freedom for students to figure out solutions on their own. I only hope a modicum of your leadership skills somehow rub off on me by the time I leave. Most importantly, thank you for your saintly patience with me and for helping me achieve my childhood dream of being a scientist.

My other committee members have also all left their mark on me and this dissertation in complementary ways. Tim Smith provided a great source of knowledge, experience, and ideas for anything having to do with optics, EP, mechanical hardware, machining, PEPL history, etc. I don't know how many times I've stopped by your office to chat about some problem, and you waltz over to the book shelf, pick out a reference, and flip right to the relevant page. Azer Yalin, thank you for asking some very smart and pointed questions and engaging with me in email discussions. You helped me clarify some of my thoughts and solve some of the early problems with the experimental setup. John Foster, thank you for providing some help and experience with the early cathode experiments. Chatting with you when I get the chance is always fun and helpful. You all provided some very thoughtful comments that I think made the dissertation markedly better after the edits.

Much of what I know about Hall thrusters, plasmas, diagnostics, and vacuum technology is thanks to the older PEPL gang in one way or another. Most of all, David and Robbie generously fielded a daily barrage of questions from a newbie about laser diagnostics and high-speed diagnostics, respectively. Thanks to Adam for showing me the ropes with Junior when I was setting up my first experiment. Ray showed me much of the equipment and processes in the lab and provided a paragon of organization that I can never hope to match. Mike McDonald saved me a lot of time by helping me work out issues with motion stages and Langmuir probes. Laura, Ricky, Rohit, and Tom all provided some help as well as a cheerful atmosphere and some laughs in the lab. Thanks also to Ray and Pete, who were patient enough to let me piggy back

with a few early tests of the TFLIF system. The movie nights and feasts (including eating contests) with some of these older PEPL students were fun moments that I will never forget.

I am convinced that Mike Sekerak, a “younger” student who started in the lab after I did but graduated earlier, possesses superhuman discipline and work ethic. What else could explain quickly finishing a PhD while also taking care of the demands of a family? Frans, you were almost the second student to arrive later and defend sooner, but I beat you to the defense by a week. Of all of my contemporaries at PEPL, I think I shared late nights in the lab most often with Roland. You added a welcome sense of insanity to the lab that will never be matched. I thoroughly enjoyed our time working together as well as the gardening, eating contests, drinking, and other lighter moments. Not to mention some mutually embarrassing episodes that I hope we will take to the grave. PEPL has been much less entertaining since you graduated. Kim spearheaded much of the repair and refurbishment of LVTF leading up to my campaign. I enjoyed traveling with Kim and Horatiu (not quite a PEPL student but a “cousin”) to IEPC in Japan. Kim is also the only PEPL student with whom I shared the trial by fire of first-year Applied Physics coursework. Thanks to all of the Applied Physics students in my cohort for helping each other through it. It was also a pleasure to work and chat with Iverson, another “cousin” frequently around PEPL. A special thanks to Iverson and Eric, who sometimes helped water the garden when I was away.

As for the younger students, firstly a big thanks to Marcel for his willingness to work absolutely as much as necessary to get the job done. With all of the trouble we had and the short time scheduled for the campaign presented in Chapter 7, the capstone of this dissertation simply would not exist without you. I may give you a hard time about not being careful enough (and sometimes you're not!), but I'm always impressed with how fast you can get work done, even with a few mistakes. Of course, we both know I may make a few mistakes myself (shh, don't tell!). To the other younger students with whom I have not worked as closely: Scott, Tim, Ethan, and Sarah, you have an amazing opportunity and are already well on your way to making your mark on the lab and producing excellent PhD work. Make the most of your time here. Trust me, it goes fast, even if you stay for an extended time.

On a more personal note, I would like to thank the people who helped keep me (somewhat) sane all these years by providing the necessary support and motivation. A weekly game night with Rich, Jason, and Bryan was at times almost the only thing keeping me going.

Concurrently with my graduate studies, and perhaps more importantly, Katie taught me a thing or two about love. Finally, I thank all of my family and especially my parents for their unwavering support and fluorescent pride as I worked my way through and completed what is, to the best of my knowledge, the first PhD in our family.

Christopher J. Durot

Ann Arbor, MI

May 2016

Table of Contents

Dedication	ii
Acknowledgements.....	iii
List of Figures	x
List of Tables	xvii
List of Appendices	xviii
List of Abbreviations	xix
Abstract.....	xx
Chapter 1 Introduction	1
1.1 Context of Study.....	1
1.2 Problem Statement	2
1.3 Aim and Scope	3
1.4 Contributions.....	4
1.5 Overview of the Dissertation	4
Chapter 2 Background	7
2.1 Spacecraft Propulsion.....	7
2.2 Electric Propulsion	8
2.3 Elementary Hall Thruster Design and Operating Principles	10
2.3.1 Magnetic Layer Hall Thrusters	10
2.3.2 Orificed Hollow Cathodes	12
2.3.3 Propellant	14
2.3.4 Hall Thruster Design Variations	14
2.4 H6 Hall Thruster.....	16
2.5 High-Speed Diagnostics.....	16
2.5.1 High-Speed Dual Langmuir Probe.....	17
2.5.2 High-Speed Imaging and the Spoke Instability	20
2.6 Hall Thruster Physics	25
2.6.1 Mode Transitions	25
2.6.2 Breathing Mode Theory and Simulation.....	32

Chapter 3 Laser-Induced Fluorescence.....	39
3.1 Principles of LIF Measurements	39
3.2 Line Broadening Mechanisms.....	41
3.2.1 Doppler Shift and Broadening	42
3.2.2 Hyperfine and Isotopic Splitting.....	44
3.2.3 Lifetime Broadening	46
3.2.4 Fluorescence Signal Saturation and Power Broadening	47
3.2.5 Lineshape of the Transition	50
3.3 Time-Resolved LIF Techniques.....	52
Chapter 4 The Theory of the TFLIF Technique	62
4.1 Definitions and Conventions.....	62
4.2 Apparatus	65
4.3 Signal Processing	66
4.3.1 Band-Pass Electronic Filter.....	68
4.3.2 Phase-Sensitive Detection.....	81
4.3.3 Fourier Analysis.....	91
4.3.4 Triggered Average	98
4.3.5 General-Purpose Benchmark	100
Chapter 5 A Practical Implementation of the TFLIF Technique	105
5.1 Hardware Considerations	106
5.1.1 Digitizer	106
5.1.2 Redundant Array of Independent Drives (RAID).....	110
5.1.3 Acousto-Optic Modulator	111
5.2 Phase Locking	114
5.2.1 Defining the Problem of Phase Locking for the TFLIF System.....	114
5.2.2 Difficulties in Finding the Optimal Reference Frequency and Phase Offset.....	116
5.2.3 A Solution to the Phase Locking Problem.....	117
5.2.4 External Versus Internal Digitizer Clock.....	122

5.2.5	Phase Unwrapping	124
5.3	Data Recycling	125
5.3.1	What is Data Recycling and Why is it Important?	126
5.3.2	A Test of Data Recycling.....	127
5.3.3	Guideline for Setting the Value of C	131
5.4	Other Software Considerations	131
5.4.1	Memory Management and Temporary Vectors	132
5.4.2	Parallelization	133
5.4.3	Chunk Length and the Power of a Power of Two.....	135
5.4.4	Setting the Time Constant of Phase-Sensitive Detection	138
5.5	Optimizing the Signal-to-Noise Ratio.....	141
5.5.1	Noise Spectral Density.....	142
5.5.2	The Raw Signal-to-Noise Ratio	145
5.5.3	Measuring the Signal-to-Noise Ratio.....	146
5.5.4	SNR vs. Laser Modulation Frequency with the Hollow Cathode	148
5.5.5	The Effect of AOM lenses and Sampling Frequency	151
5.5.6	SNR vs. Laser Modulation Frequency with the H6.....	152
5.5.7	Electronic Band-Pass Filter.....	154
Chapter 6	Validation Experiments for TFLIF with a Hollow Cathode	156
6.1	Facility and Plasma Source	157
6.2	Preliminary Validation Considerations	159
6.3	Validating the TFLIF system with Sinusoidal Discharge Current Oscillations.....	160
6.3.1	Motivation for an Experiment with Sinusoidal Discharge Oscillations	160
6.3.2	Results.....	161
6.4	Validating the TFLIF system with Random Discharge Current Oscillations	164
6.4.1	Motivation for an Experiment with Random Discharge Oscillations.....	165
6.4.2	Generating Random Oscillations	166
6.4.3	The Distribution of the Random Variable that Generates the Frequency List	166

6.4.4	Consistency Checks on the Properties of the Simulated Triggered Average	168
6.4.5	Advantages and Limitations of List Mode.....	171
6.4.6	Results.....	172
6.5	Summary of Arguments Validating TFLIF.....	175
Chapter 7 Initial Hall Thruster Campaign		177
7.1	Experimental Configuration.....	178
7.1.1	Experimental Setup in the Vacuum Chamber.....	178
7.1.2	Auxiliary Measurements	181
7.1.3	Test Matrix.....	182
7.2	Results and Discussion.....	184
7.2.1	Evolution of the Velocity Distribution Between Two Axial Positions.....	184
7.2.2	The Effect of Varying Magnetic Field.....	187
7.2.3	Validation of TFLIF.....	194
Chapter 8 Discussion and Conclusion		205
8.1	Comparison of Features and Specifications of Time-Resolved LIF Techniques.....	205
8.1.1	Acquisition Time per Laser Wavelength	207
8.1.2	Signal Bandwidth.....	208
8.1.3	Post-Processing.....	209
8.1.4	Monetary Cost.....	209
8.2	The Development of TFLIF	209
8.3	Hall Thruster Physics	211
8.4	Future Research.....	212
8.5	Conclusion.....	214
Appendices.....		215
References.....		225

List of Figures

<p>Figure 2.1. Cross-sectional diagram of a basic, magnetic layer Hall thruster. The key parts in construction, electric and magnetic fields, and representative paths of ions, electrons and neutrals are shown. Reproduced from Figure 7-1 in [12].</p>	11
<p>Figure 2.2. Cross-sectional diagram of an orificed hollow cathode. Image reproduced from Figure 6-4 in [12].</p>	13
<p>Figure 2.3. Heat maps showing how the electron density changes over one period of a breathing oscillation of the BHT-600. The data are synthesized as the output of the transfer functions at each spatial point for a single input discharge current trace. Reproduced from [36].</p>	19
<p>Figure 2.4. The synthesized characteristic electron density from the average transfer function (red) reproduces the key features of the original signal (black). Reproduced from [37].</p>	20
<p>Figure 2.5. The 2D array of a single frame of the video is azimuthally binned and the intensity is averaged over 180 bins to generate a 1D array of bin intensity as a function of azimuthal angle. The complete video is then represented by a 2D array of dimensions $180 \times N_f$. Reproduced from Ref. [40].</p>	22
<p>Figure 2.6. Top: example 2D power spectral density of a $180 \times N_f$ array of azimuthal intensity as a function of time. Bottom: vertical cross-sections of the 2D power spectral density are taken for a number of spoke modes with $m > 0$ and the $m = 0$ mode. Reproduced from Figure 6 of Ref. [6].</p>	23
<p>Figure 2.7. Left: Plot of mean discharge current versus magnetic field magnitude over the six mode regions. Right: High-frequency oscillation amplitude plotted as a function of magnetic field magnitude for frequencies of (1) 20; (2) 60; (3) 140 MHz at a floating probes just downstream of the exit plane. Reproduced from figures 5 and 4d in [42].</p>	26
<p>Figure 2.8. Photos of the H6 plume (left) and contours of constant intensity (right) for both low-current mode (upper) and high-current mode (lower) [43]. Reproduced from Figure 1.</p>	28
<p>Figure 2.9. Examples of magnetic field sweeps showing transitions in mean discharge current (solid lines) and RMS oscillation amplitude (dotted lines) as a function of magnetic field magnitude from Sekerak et al. [10]. Reproduced from Figure 20 (a).</p>	29
<p>Figure 2.10. Examples of the standing wave in ion density (left, in units of 10^{12} cm^{-3}) and slight oscillation in the position of the electric field (right in units of V) from the numerical model of Boeuf and Garrigues [52]. Reproduced from Figures 9 and 10.</p>	35
<p>Figure 2.11. Simulations show that Joule heating is balanced mostly by convection in the breathing mode (a) but by wall loss in the stable mode (b) [55]. Reproduced from Figure 11.</p>	37
<p>Figure 3.1. Partial Grotrian diagram of Xe II energy levels and relatively strong transitions between them. The laser transition at 834.7 nm is highlighted in red and the fluorescence transition at 541.9 nm is highlighted in green. Note energy is in wavenumbers relative to the Xe II ground state. Modified from FIG. 2 of reference [60].</p>	40

Figure 3.2. A simulation of the power broadening of the Xe II 5d4D7/2 – 6p4P5/20 transition for a variety of laser power levels. The first profile where the side peak is noticeably higher corresponds to $\max I0/Isatv = 0.79$, and the first profile with a significantly broadened main peak is at $\max I0/Isatv = 7.92$. Reproduced from [62]. 50

Figure 3.3. Left: A slightly Doppler broadened lineshape of the 5d2F7/2 – 6p2D5/2o transition has a FWHM of 750 MHz (reproduced from [67]). Right: the Doppler-free lineshape measured in another study has a FWHM of 440 MHz (reproduced from [65]). The widths are only 626 m/s and 367 m/s in speed. . 51

Figure 3.4. Ion VDF measured 4 mm downstream of the exit plane of a 200-W permanent magnet Hall thruster [84]. Reproduced from Fig 5. 56

Figure 3.5. The time-resolved ion VDF at the exit plane of a cylindrical Hall thruster, reconstructed to second order in the harmonic decomposition [11]. Reproduced from FIG. 4. 58

Figure 3.6. Axial time-resolved ion VDF of the BHT-600 at 15 mm downstream of the exit plane on the thruster centerline [87]. Reproduced from Figure 8..... 61

Figure 4.1. Diagram of the experimental setup showing a block diagram of airside optics sans mirrors, a schematic of the optics and hollow cathode plasma source inside the chamber, and a block diagram of instruments post collection outside the chamber. 65

Figure 4.2. Flowchart of the post-processing steps on the PMT signal to recover the TRLIF signal. Though not explicitly shown in the flowchart, note that a high-speed signal related to the LIF signal (e.g. discharge current or floating probe) is necessary either to form transfer function estimators or to define the triggers used for triggered averaging. 67

Figure 4.3. Bode plot of the frequency response of the series high-pass/low-pass filter with cutoff frequencies $f_{mod} \pm 0.1 f_{mod}$ for $f_{mod} = 1$ MHz. 70

Figure 4.4. Amplitude gain profiles of band-pass filters centered on 1 MHz for three representative bandwidths from 0 to 1 MHz. 74

Figure 4.5. The SNR improvement factor calculated by solving Equation (4.8) for $f_{mod} = 1$ MHz and varying the bandwidth of a band-pass filter centered on f_{mod} . The calculation, assuming white noise, shows that the SNR improvement monotonically increases but converges to a value of about 5.5 as the bandwidth approaches zero. 75

Figure 4.6. An illustration showing that signal modulation of the envelope with a cosine waveform shifts the signal bandwidth from centered on zero to two sidebands centered on the modulation frequency f_{mod} with the same bandwidth as the double-sided bandwidth of signal envelope $F(t)$ 77

Figure 4.7. Gain and phase shift from the Bode plot of Figure 4.3, zoomed in on the bandwidth interval $[f_{mod} - 20$ kHz, $f_{mod} + 20$ kHz]. 78

Figure 4.8. The output of phase-sensitive detection $V_{out}(t)$ given by solutions to Equation (4.31) for a few τ values, which show the increasing response time and decreasing ripple as τ increases. 86

Figure 4.9. Gain of the $2\omega_{mod}$ ripple for 1-MHz and 2-MHz modulation frequencies (blue dashed and dotted lines) plotted together with the maximum frequency with gain exceeding $G_{min} = 0.95$. Both are functions of the time constant τ . The ripple gain determines the minimum acceptable τ while the maximum frequency passed determines the maximum acceptable τ 90

Figure 5.1. The SNR moderately increases as a function of R_{sat} 109

Figure 5.2. An Illustration showing the method of finding ϕ_{eff} at a single point in time. PSD output at ten different reference phase offsets (blue circles) are used in the least squares fit to the function $\sin(\phi_{ref} + \phi_{jit})$ (red line). The maximum occurs at $\phi_{ref} = \phi_{eff} = \pi/2 - \phi_{jit}$ (open blue circle). 120

Figure 5.3. The effective phase offset at all points (green line) is interpolated from a series of effective phase offset points measured throughout the 60-s dataset (open blue circles). Top: example of the "best case" scenario with negligible drift and low noise. Bottom: A difficult example with fast, nonlinear drift in the effective phase offset. 121

Figure 5.4. The internal digitizer clock can result in fast drift in the difference between the reference and signal phases (green line) that would break the algorithm in Subsection 5.2.3. An intermediate step fitting the modulation frequency is necessary in this case..... 124

Figure 5.5. Top: A conceptual illustration showing how windowing consecutive chunks leads to large sections of data that do not contribute significantly to the DFT. Bottom: An example illustrating data recycling with $C = 2$, with the first cycle chunks shown in blue and the second cycle chunks shown in red. N_{chunks} is the total number of chunks taken from the dataset. 126

Figure 5.6 The measured RMS residual between TFLIF and TALIF signals as a function of the number of cycles in the transfer function average (blue diamond). Also shown is the expected RMS residual following the A/C function expected for averaging over the same number of chunks if they were independent. The measured noise reduction closely follows the independent averaging trend only for $C = 2$, and lags increasingly further behind thereafter. 129

Figure 5.7. TFLIF signal for $C = 1$ (top), $C = 5$ (upper middle), and $C = 10$ (lower middle) converge toward the TALIF signal (bottom) as the number of cycles C increases. 130

Figure 5.8. Plot of the average computational time of the FFT for all time series lengths between 2^{14} and 2^{15} . Even lengths (blue dot) are generally faster than odd lengths (red dot). Moreover, prime lengths (red circle) perform worst and powers of two (blue circle) perform best. 137

Figure 5.9. Raising the time constant from $1 \mu s$ to $8 \mu s$ markedly improves SNR while producing barely noticeable distortion of the waveform. 139

Figure 5.10. Plots of the difference in TFLIF signal waveform analyzed with $\tau = 2, 4,$ and $8 \mu s$ from the waveform with $\tau = 1 \mu s$. The systematic differences for $\tau = 4 \mu s$ and $\tau = 8 \mu s$ imply waveform distortion, while distortion present for $\tau = 2 \mu s$ is below the noise. 140

Figure 5.11. Measured noise spectral density of the PMT signal for the H6 operating at 150-V discharge voltage, 10-mg/s anode flow rate. The nonwhite spectra suggest background collected light is the dominant noise source. The measured noise spectra are all orders of magnitude above the nominal amplifier input noise or calculated resistor Johnson noise (shown in blue)..... 143

Figure 5.12. Signal-to-noise ratio of time-averaged LIF signal reaches a maximum near a laser modulation frequency of 500 kHz and falls toward zero thereafter. Filtering about 50% of collected light dramatically reduces SNR by more than 50% at high modulation frequencies, but only by 20% at most at the low modulation frequencies.	149
Figure 5.13. AOM lenses with a 120-mm focal length allow the SNR to reach a maximum near a modulation frequency of 1 MHz, significantly higher than with the 200-mm lenses of Figure 5.12. The SNR is insensitive to the sampling frequency at these conditions.	152
Figure 5.14. The SNR with the H6 reaches a maximum at the laser modulation frequency of 2MHz. The SNR is much lower at the point 15 mm downstream of the exit plane than at 4 mm upstream, but the maximum occurs at the same frequency for both.	154
Figure 5.15. SNR as a function of electronic filter bandwidth with $f_{mod} = 100$ kHz. SNR does not depend strongly on the electronic filter bandwidth.	155
Figure 6.1. Photographs of the "Junior" Chamber at PEPL (left) and the cathode used in the validation campaign (right).	157
Figure 6.2. Diagram of the optics and plasma source used in the validation campaign. Laser light follows the red path, collected fluorescence light follows the green path, and the connections of the power amplifiers to the anodes and function generator are shown in yellow.	158
Figure 6.3. The discharge current oscillated between about 4.25 A and 6.1 A with a sinusoidal oscillation fixed at 10 kHz.	160
Figure 6.4. A comparison of the average of time-resolved VDF profiles from transfer function averaging (red "+") and triggered averaging (green "x") shown against the average VDF profile measured by the lock-in amplifier LIF system (blue line).	162
Figure 6.5. Heat maps of the TFLIF (top) and TALIF (middle) signals and their residual (bottom) as functions of velocity and time show no systematic differences between the two signals.	163
Figure 6.6. "Snapshots" of the VDF profile in time for the transfer function average (red "+") and triggered average (green "x"). The solid lines are spline interpolations meant to be a guide to the eye without a precise physical meaning.	164
Figure 6.7. The discharge current oscillated between about 5.5 A and 8.5 A with a sinusoidal oscillation with a randomly varying period in the range of 7.5 - 12.5 kHz.	165
Figure 6.8. Plot of simulated triggered averages for $\Delta f = 3$ kHz (blue) and $\Delta f = 5$ kHz (green) together with the measured triggered average of TRLIF signal from a preliminary test with the X2. Least squares fits to the functional form in Equation (6.1) are also plotted in dotted lines.	168
Figure 6.9. Plot of the phase difference of 1000 instances of the random waveform relative to the phase of at sinusoid at a frequency f_0 (blue lines) for the $\Delta f = 5$ kHz case. The linear fit representing the average phase drift of the random waveforms from the sinusoid at f_0 is also plotted in red.	170

Figure 6.10. A Comparison between TRLIF signal from the transfer function average (top) and triggered average (middle) with the residual on the bottom. The input to the transfer function is an actual segment of discharge current beginning at a trigger so that they are initially in phase..... 173

Figure 6.11. A Comparison between LIF signal from the transfer function average (top) and triggered average (middle) with the residual on the bottom. The input to the transfer function is the triggered average of discharge current traces. In this case, the TFLIF signal does reproduce the TALIF signal as expected. 174

Figure 7.1. Diagram of the experimental setup inside the Large Vacuum Test Facility. Reproduced from Ref. [101]. 179

Figure 7.2. The H6, alignment pin, floating probe, and collection optics photographed in position for the campaign. 180

Figure 7.3. Photograph of the warped outer pole piece near the 11 O'clock position of the H6..... 180

Figure 7.4. Calibration curves for the cathode (left) and anode (right) flow controllers. 182

Figure 7.5. Example of flow rate measurement drift observed after equilibrium in back pressure at a constant flow rate setting of 50 sccm on the anode controller. 182

Figure 7.6. The Ion velocity distribution at $z = -4$ mm for the 150-V, 10-mg/s operating condition of the H6 with nominal magnetic field shows an oscillation in central velocity, distribution spread, and relative metastable density. 184

Figure 7.7. Ion velocity distribution as a function of time at $z = 15$ mm for the 150-V, 10-mg/s operating condition of the H6. 185

Figure 7.8. Ratio of measured ion velocity distribution FWHM at $z = 15$ mm to that predicted by bunching given the ion velocity distribution at $z = -4$ mm. It oscillates between approximately 1 and 3, highly correlated with discharge current oscillations, and possibly indicates axial motion of the ionization zone. 186

Figure 7.9. Power spectral density of FASTCAM intensity for the $m = 0$ breathing mode and spoke modes $m = 1$ through $m = 3$ for the 150-V, 10-mg/s operating condition of the H6 at three magnetic field settings. The thruster exhibits an extremely strong $m = 0$ mode at $Br/Br^* = 0.52$ (left). Some azimuthal perturbations visually appear at the nominal condition $Br/Br^* = 1$ (middle). Azimuthal perturbations are stronger but still dominated by the global oscillation at $Br/Br^* = 1.48$ (right)..... 188

Figure 7.10. The spoke surface at low-field condition (left) exhibits a strong global mode oscillation. Azimuthal perturbations begin to appear by the nominal setting (center) and intensify at the high-field setting (right)..... 189

Figure 7.11. Discharge current power spectra for the low-field setting with $Br/Br^* = 0.52$ (left), the nominal $Br/Br^* = 1$ setting (center), and the high-field setting with $Br/Br^* = 1.48$ (right). The current spectra are similar to the $m = 0$ mode but also captures higher-frequency features. 190

Figure 7.12. The mean discharge current (squares connected with solid line) is approximately constant while the oscillation amplitude, measured by RMS (dashed lines) or extrema (dotted lines), changes drastically during the magnetic field sweep..... 191

Figure 7.13. Ion velocity distribution as a function of time at $z = 0$ mm for the 150-V, 10-mg/s operating condition of the H6 with $B_r/B_r^* = 0.52$. Note that the metastable ion population probed appears to collapse almost entirely to the noise floor after the distribution reaches a maximum in mean velocity (visible as vertical stripes of mostly dark blue color). 193

Figure 7.14. Ion velocity distribution as a function of time at $z = 0$ mm for the 150-V, 10-mg/s operating condition of the H6 with $B_r/B_r^* = 1.48$. The mean velocity oscillates nonperiodically and within a smaller range than it does at the lower field conditions..... 193

Figure 7.15. A comparison showing the time-averaged ion VDF from TFLIF agrees with a traditional lock-in amplifier measurement of the time-averaged VDF. This example is for the $z = 0$ mm, $B_r/B_r^* = 1.48$ data set..... 195

Figure 7.16. A comparison of time-averaged LIF profiles for the $z = 0$ mm, $B_r/B_r^* = 0.52$ dataset. This dataset gave the worst agreement of the points in this paper, but the time-averaged ion VDF from TFLIF still generally agrees with a traditional lock-in amplifier measurement of the time-averaged VDF. The relatively poorer agreement is likely due to trouble with laser tuning while acquiring this point. 196

Figure 7.17. The TFLIF and triggered average signals are nearly identical at the low-field setting, as expected if the assumptions of both averaging techniques are satisfied. 197

Figure 7.18. A direct comparison between the TFLIF signal and the triggered average LIF signal for the $z = 0$ mm, $B_r/B_r^* = 1.48$ dataset shows significant systematic differences. The triggered average shows a chaotic oscillation in the mean velocity but the triggered average fails to detect nonperiodic oscillations, hence another benchmark is required for nonperiodic oscillations..... 198

Figure 7.19. An example of the general-purpose benchmark with the $z = 0$ mm, $B_r/B_r^* = 1.48$ data set. The transfer function faithfully reproduces the triggered average LIF signal when the triggered average discharge current is used as input to the transfer function. 199

Figure 7.20. A direct comparison of TFLIF signals from discharge current (top) and floating probe voltage (center) input signals for the low-field setting $B_r/B_r^* = 0.52$ shows small systematic differences in the residual (bottom). 201

Figure 7.21. The TFLIF signals from the discharge current (top) and the floating probe voltage (center) input signals capture the same general features for the high-field setting $B_r/B_r^* = 1.48$. The residual (bottom) is almost entirely random noise in the second half of the window, while the first half of the window shows differences in the amplitude of the mean velocity oscillation..... 202

Figure 7.22. The mean velocity (top) and the relative metastable ion density (bottom) of the TFLIF signal from the discharge current (blue) and floating probe (green) inputs agree for most of the window except the first 0.1 ms. These VDF properties were calculated from the dataset of Figure 7.21..... 203

Figure 7.23. The input signals used to generate the TFLIF signals in Figure 7.21 are completely different, but the TFLIF signals output from their respective transfer functions are very similar. 204

Figure A.1. An illustration of the shape of the Tukey window for a selection of r values and a time series length of 1000. 216

Figure A.2. Heat maps of the TFLIF signal (top), TALIF signal (middle), and their residual (bottom) for four parameter values of (A) $r = 0$, (B) $r = 0.1$, (C) $r = 0.25$, and (D) $r = 0.5$. Both random and systematic error are smaller for large r values. 220

List of Tables

Table 2.1. A summary of the quantitative criteria for defining H6 mode transitions in terms of FASTCAM power spectrum peaks and discharge current magnitude and oscillation amplitude developed by Sekerak [44]. Reproduced from Table 4.2.....	30
Table 3.1. Xenon isotopic abundance in percent from reference [28].	45
Table 4.1. Select values of the minimum allowable ratio of time constant to modulation period $\min(\tau/T_{mod})$ given values of $G_{2\omega}$, a threshold of the maximum ripple amplitude to DC amplitude in the phase-sensitive detection signal.	89
Table 5.1. A summary of the operating characteristics of two sets of focusing/collimating lenses. There is a trade-off between the maximum frequency f_{max} and the diffraction efficiency η_{diff}	113
Table 7.1. Test matrix of conditions where TFLIF signal was captured. The campaign was entirely at the H6 150-V, 10-mg/s operating condition with all LIF data taken along the channel centerline at 3 O'clock. Green cells denote data was taken and presented. Analysis failed to yield reasonable results for data taken at the exit plane with nominal field (red cell) for unknown reasons.....	183
Table 8.1. Summary of the five most relevant features of four TRLIF techniques used with Hall thrusters or similar EP devices. The most attractive features of each category are highlighted in green. Tilde denotes an estimate that is likely accurate within a factor of about 3 or better.	206
Table A.1. The RMS residual between TFLIF and TALIF signals decreases as the r parameter of the Tukey window increases.....	218

List of Appendices

Appendix A: Window Effects.....	216
Appendix B: An Analytical Model of Synchronized Photon Counting.....	221

List of Abbreviations

AOM	acousto-optic modulator
CW	continuous wave
DFT	discrete Fourier transform
ENBW	equivalent noise bandwidth
ETFE	empirical transfer function estimator
FFT	fast Fourier transform
HDLP	high-speed dual Langmuir probe
FWHM	full width at half maximum
LIF	laser-induced fluorescence
LP	Langmuir probe
PEPL	Plasmadynamics and Electric Propulsion Laboratory
PMT	photomultiplier tube
PSD	phase-sensitive detection
RAID	redundant array of independent disks
RMS	root mean square
SNR	signal-to-noise ratio
TALIF	Triggered Average LIF
TFLIF	Transfer Function LIF
TRLIF	time-resolved LIF
VDF	velocity distribution function

Abstract

Development of a Time-Resolved Laser-Induced Fluorescence Technique for Nonperiodic Oscillations

by

Christopher J. Durot

Chair: Alec D. Gallimore

Time-resolved measurements of ion dynamics could be key to understanding the physics of instabilities, electron transport, and erosion in Hall thrusters. Traditional measurements of the ion velocity distribution in Hall thrusters using laser-induced fluorescence (LIF) are time-averaged since lock-in amplifiers must average over a long time constant for a reasonable signal-to-noise ratio. Over about the past decade, at least four other time-resolved LIF techniques have been developed and applied to Hall thrusters or similar plasma devices. One limitation of these techniques is the implicit assumption of periodic oscillations in the averaging scheme. There is a need for a more general technique since Hall thrusters can operate with nonperiodic oscillations that vary unpredictably. This dissertation presents the development of a time-resolved LIF (TRLIF) technique that addresses this need. This system averages the signal using a combination of electronic filtering, phase-sensitive detection, and Fourier analysis. A transfer function is measured to map an input signal (such as discharge current) to an output signal (TRLIF signal). The implicit assumption of this technique is that the input is related to the output by a time-invariant linear system, a more general assumption than periodicity. The system was validated using a hollow cathode with both periodic and random discharge current oscillations. A series of benchmark tests was developed to validate the signal by verifying that it satisfies theoretical expectations. The first campaign with the H6 Hall thruster demonstrated signal recovery in both periodic and nonperiodic modes. Measurements of the evolution of the ion flow downstream show that kinematic compression explains the width of the ion velocity distribution only at certain phases of the oscillation. A distinct change in ion dynamics was detected as the magnetic field magnitude increased: a high-amplitude, relatively periodic oscillation in the ion velocity distribution gave way to a low-amplitude, chaotic oscillation. High amplitude oscillations of the mean ion velocity suggest that the bimodal distributions detected at many operating conditions (with time-averaged measurements) are the result of oscillations.

Chapter 1

Introduction

The journey of a thousand miles begins with one step.

– Lao Tsu

1.1 Context of Study

Although Hall thrusters have a conceptually simple design and have been in development since the 1960s with abundant research in the US [1] and former USSR [2], among others, one of the main threads in Hall thruster research to this day is to understand the fundamental plasma physics at work in a thruster. One of the most important topics of research in Hall thruster physics is to understand the nature of oscillations and instabilities and how they affect thruster performance [3], [4]. Another important topic of research is to identify the sources of the anomalous electron transport across the magnetic field to the anode; the effective electron mobility is found to be orders of magnitude higher than the classical collisional cross-field mobility (e.g. [5], [6]). These two topics overlap significantly since there is evidence that some oscillations can affect electron mobility and can be caused by insufficient mobility to sustain the discharge. Clearly, progress in these topics depends in part on measurements of oscillations of plasma properties in Hall thrusters.

To that end, several high speed plasma diagnostic techniques have been developed in recent years to resolve oscillations in Hall thrusters to a bandwidth of about 100 kHz or more, which contains two of the largest amplitude instabilities: breathing and spoke oscillations [4]. Each technique yields different plasma properties, and together they can give complementary information that leads to a more complete picture of thruster oscillations. The high speed dual Langmuir probe technique has allowed observations of oscillations in electron density, electron temperature, and plasma potential [7]. Specialized signal processing of the video from fast framing cameras have shown oscillations in light emission intensity [8], which is related to discharge current density [9]. Time-resolved laser-induced fluorescence measurements have

given a glimpse of the dynamics of the ion velocity distribution and the electric field. Due to limitations in the averaging techniques currently in use, these measurements have been made only where natural oscillations are quasiperiodic or where an external perturbation drives periodic oscillations with operating conditions that would normally give rise to nonperiodic oscillations. The place of this dissertation in the field of research above is to develop, validate, and demonstrate a new time-resolved laser-induced fluorescence (TRLIF) technique that can be used to interrogate Hall thrusters in all normal operating conditions regardless of the nature of the oscillations.

1.2 Problem Statement

The main difficulty of LIF measurements in plasma sources such as Hall thrusters is that there is a large proportion of background light that is necessarily collected with the LIF signal light. The background light in this case comes primarily from fluorescence from electron impact excitation of the same transition for which the laser induces fluorescence. All LIF measurement techniques perform averaging in some sense over data taken during an interval in time to recover the signal from the noise. Since the signal-to-noise ratio of the raw data is far less than unity (see Subsection 5.5.2), truly time-resolved measurements are impossible to make in real time, and it is difficult to investigate short time scales.

Though there are techniques previously developed to measure time-resolved LIF signal and some of those techniques has been demonstrated with Hall thrusters (see Section 3.3), all techniques used thus far share an important disadvantage for Hall thrusters. They all work based on the assumption that the physical process being measured is at least quasiperiodic and therefore many repeated cycles of the process can be averaged in some sense to recover the signal from the noise. This is a significant disadvantage for interrogating Hall thrusters because the amplitude and frequency of breathing and spoke oscillations can vary unpredictably (cf. the broad power spectral density plots and the turbulent discharge current and spoke surfaces in Refs. [6], [10]). Thermal issues can cause measurements to drift over time, which can be problematic for a method that assumes a repeatable oscillation while averaging over a time scale comparable with the drift time scale [11]. The oscillations of Hall thrusters may have a well-defined spectral density, but they are often not periodically repeating. Thus, many studies of TRLIF in Hall thrusters to date have used a perturbation to the operating condition to force periodic oscillations that are amenable to the averaging techniques in use (see Section 3.3).

The problem that this dissertation addresses is the need for a technique to measure TRLIF in plasma sources with nonperiodic oscillations. This feature would give TRLIF measurements the same general applicability that current high-speed Langmuir probe and high-speed camera diagnostics have to make measurements on normal Hall thruster operating conditions without forcing periodic oscillations. Such a general technique may also enable similar measurements in other plasma sources with nonperiodic oscillations.

1.3 Aim and Scope

The aim of this dissertation is to investigate how this problem may be solved with a new TRLIF measurement technique. The goal is to achieve a bandwidth of about 100 kHz so that breathing and spoke oscillations are well within the bandwidth. The initial validation experiments are limited to a discharge between a hollow cathode and anode because it allows flexibility in driving the discharge current oscillation, enabling testing the new technique with a precisely periodic or truly random oscillation. An initial campaign with the H6 demonstrates that the new technique can be used to gain useful information with a Hall thruster. The study is restricted to the Xe II (singly charged xenon) $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition commonly used for Hall thruster experiments, but in principle the technique can apply to any level scheme appropriate for LIF as long as the upper level lifetime is much shorter than the desired time resolution.

An important limitation to the scope is the sense in which a measurement can be made. Noise completely obscures the signal (see Subsection 5.5.2) and a large amount of raw data must be averaged to recover the signal from noise. As already mentioned, a real-time, direct measurement of TRLIF signal is a hopeless task. The sense in which the measurement of this technique (and the others) is "time-resolved" is restricted to a characteristic signal that is calculated based on some assumptions about the nature of the signal and an averaging scheme that takes advantage of those assumptions. For example, the previous TRLIF techniques assume periodic oscillations (see Section 3.3). The new technique presented in this dissertation assumes that a time-invariant linear system relates some input signal (e.g. discharge current or floating probe potential) to the output of TRLIF signal. In that case, there is a transfer function that can be measured that maps the input signal to the output signal. It is important to keep in mind that all TRLIF measurements are indirect in the sense that they all average in some way over a large

quantity of data that includes many thousands of different periods of the oscillation, but the averaging scheme reduces the large dataset down to a single waveform only a few milliseconds long at most.

1.4 Contributions

This dissertation adds two main contributions to the field. Firstly, a technique to measure TRLIF signal in plasmas with nonperiodic oscillations is developed and demonstrated. The system is analyzed theoretically and a number of tests are performed to help determine optimal analysis parameters. The measurements are validated using both periodic and random discharge oscillations of a hollow cathode. A series of benchmarks are proposed and used to help validate the characteristic TRLIF signal. Finally, the new technique is compared with other TRLIF techniques in terms of key features such as data acquisition time, signal bandwidth, and the required financial investment.

Secondly, the TRLIF technique developed at the Plasmadynamics and Electric Propulsion Laboratory (PEPL) was applied in an initial campaign with the H6 Hall thruster at the 150-V, 10-mg/s operating condition. Several new results have been obtained that will help guide a future TRLIF campaign on the H6. Ion dynamics are strongly dependent on magnetic field magnitude, with high-amplitude, periodic oscillations at low field and low-amplitude, chaotic oscillations at high field. An analysis of the axial evolution of the ion VDF with nominal magnetic field indicates that kinematic compression explains the distribution spread at certain phases of the oscillation but not at others. This implies that other effects intermittently widen the VDF such as ionization or collisions between the two points probed.

1.5 Overview of the Dissertation

This dissertation is organized as follows. Chapters 2 and 3 contain the background material, while Chapters 4 through 8 constitute the main contributions of this dissertation to research in electric propulsion.

Chapter 2 expands on the background and context necessary to appreciate this research. It covers the principles of electric propulsion, the basic design and operation of Hall thrusters, the current state of relevant research in Hall thruster physics, and the high speed diagnostics used to interrogate Hall thrusters.

Chapter 3 provides the necessary background on laser-induced fluorescence. It first covers the principles of how LIF measurements are made and the main challenges commonly faced. The line broadening mechanisms that both enable and hinder velocity measurements are described next. The chapter concludes with a description of other time-resolved LIF techniques recently developed for use with Hall thrusters and an argument that there is a need for a new technique.

Chapter 4 describes the PEPL TRLIF scheme, including the physical apparatus and the theory behind the signal processing. The main signal processing technique is known as Transfer Function Laser-Induced Fluorescence (TFLIF). A secondary technique used for benchmarking is called Triggered Average Laser-Induced Fluorescence (TALIF). Finally, the chapter concludes with a proposed general-purpose benchmark that can be used to validate signals synthesized using the transfer function.

Chapter 5 delves into three general categories of details to making a practical implementation of a TFLIF measurement in the laboratory. The first section covers the specialized hardware considerations concerning the digitizer, AOM, RAID, and multichannel function generator that determine how those parts are chosen and the limits that they impose. Secondly, there are many software considerations such as modulation phase locking, data recycling, memory management, and parallelization to consider. Ultimately, thousands of lines of MATLAB code are used in the data analysis and presentation of results. The final section presents preliminary experiments that guided the choice of parameters and algorithms to maximize the signal-to-noise ratio (SNR).

Chapter 6 presents the basic validation experiments of the PEPL TFLIF system using both periodic and random oscillations in a hollow cathode discharge with an external anode. The chapter demonstrates a number of benchmarks as unambiguously as possible since the oscillations are explicitly controlled. Section 6.4 demonstrates that the TFLIF technique is capable of recovering chaotic or even random oscillations, a key result necessary for truly time-resolved measurements at some Hall thruster operating conditions.

Chapter 7 presents the first campaign in a Hall thruster using the TFLIF technique. It demonstrates that a practical campaign with a Hall thruster is possible and gives interesting new results on the ion dynamics in the H6 Hall thruster at the 150-V, 10-mg/s operating condition.

Chapter 8 concludes the dissertation with three discussions. First, it compares the advantages and disadvantages of many of the TRLIF techniques that have been used in electric propulsion, including the technique developed in this dissertation. Second, the main results and their implications for both the development of the technique in general and Hall thruster physics are discussed. Finally, potential avenues for future research are listed before closing the dissertation.

Chapter 2

Background

If I have seen further than others, it is by standing upon the shoulders of giants.

– Isaac Newton (Letter to Robert Hooke, 15 February, 1676)

This chapter provides the necessary background regarding electric propulsion (EP). A complete survey of EP is beyond the scope of this dissertation. The first two sections are meant only to provide some context and motivation for the study of EP. They set the stage for the following two sections, which provide an introduction to Hall thrusters (Section 2.3) and introduce the Hall thruster used in this dissertation (Section 2.4). The final two sections review high-speed diagnostics used with Hall thrusters (Section 2.5) and review recent research on the physics of Hall thrusters relevant to this dissertation (Section 2.6).

2.1 Spacecraft Propulsion

The principle of conservation of momentum requires that any spacecraft maneuver changes not only the momentum of the spacecraft but also the momentum of something else, in general called a reaction mass. This presents a challenge for spacecraft since in general they are isolated from other objects. In the absence of other objects to push, the standard solution to this problem is the rocket, which carries its reaction mass (propellant) onboard to expel in the direction opposite the desired thrust. Solar sails, space tethers, and the more exotic Bussard ramjet are examples of some concepts that cleverly manipulate some external reaction mass to generate thrust and therefore require no propellant to be carried. Rockets are nevertheless the mainstay space propulsion system for the foreseeable future due to technological limitations and the fundamental limits on the availability of the naturally occurring reaction mass that such concepts exploit. It is therefore critical to the future of space exploration to understand and push the limits of the capabilities of rockets.

Spacecraft missions are often characterized by the velocity increment necessary for the mission (Δv). The Tsiolkovsky rocket equation provides a basic understanding of how the

velocity increment imparted by a rocket is related to mass delivered (m_d), propellant mass used (m_p), and the exhaust velocity (v_{ex}). It can be derived using the conservation of momentum [12]. It is valid given certain simplifying assumptions such as a single stage rocket under no force of gravity and that all mass is exhausted at a single, constant velocity with no plume divergence. One form of the rocket equation is:

$$\frac{m_d + m_p}{m_d} = e^{\Delta v/v_{ex}}. \quad (2.1)$$

The key takeaway is that the propellant required is an exponential function of the ratio of velocity increment to exhaust velocity ($\Delta v/v_{ex}$). Therefore, exhaust velocity should be at least close to the magnitude of Δv , otherwise the rocket may only reach the desired velocity increment with an absurdly high mass ratio, the ratio of the delivered mass plus propellant to the delivered mass (left hand side of Equation (2.1)). It follows that the exhaust velocity is a key figure of merit for rocket performance. It is normally quoted in terms of specific impulse I_{sp} , which is defined to be the thrust per rate of propellant use by weight on Earth:

$$I_{sp} = \frac{\dot{m}v_{ex}}{\dot{m}g} = \frac{v_{ex}}{g}. \quad (2.2)$$

2.2 Electric Propulsion

It is not enough to have propellant mass to eject; all rockets have a source of energy that is converted into the kinetic energy of the spacecraft as well as losses such as heat and the kinetic energy of the propellant. The energy for chemical rockets is stored within the propellant itself and is released during a chemical reaction. The exhaust velocity of chemical rockets is fundamentally limited, as Jahn notes [13], by (1) the energy available in the chemical reaction, (2) the heat load on the combustion chamber or nozzle, and (3) losses to internal modes of the gas or radiation. Due to these limitations, the exhaust velocity for chemical rockets is at most about 5000 m/s (see Table 1-2 in Ref. [13] and Table 1-1 in Ref. [12]). This is poorly matched to the velocity increment necessary for many interplanetary missions from Earth orbit, often on the order of 10^4 - 10^5 m/s (see Table 1.1 in Ref. [13]).

Electric propulsion (EP) systems have one common feature: they use electrical energy to accelerate the propellant. There is a wide variety of electric propulsion systems in terms of the physical mechanisms they exploit and the propellants used [14], and a correspondingly wide

variety of features such as thrust, I_{sp} , power, weight and lifetime [15]. There is typically some form of plasma discharge, and thrust is generated primarily by ion acceleration since they are much more massive than electrons. There are three main forms of EP distinguished by the form of propellant acceleration used: (1) electrothermal propulsion uses an electrical current to heat the propellant, (2) electrostatic propulsion accelerates ions using an electric field, and (3) electromagnetic propulsion accelerates ions through an interaction of an electrical current within the plasma with a magnetic field.

Typical EP systems have an exhaust speed of several tens of kilometers per second (see table 1-1 in Ref. [12]), well above the limits of chemical rockets. It can exceed the limits of chemical rockets because (1) the source of energy is decoupled from the propellant and limited by the power supply available, not the intrinsic energy density of the propellant, and (2) the heat load on the thruster is tolerable since the plasma is low density.

As discussed above, higher exhaust speed dramatically reduces the required mass of propellant, which leads to a number of potential advantages. EP can enable missions that would otherwise be impractical or cost-prohibitive. For example, EP enabled the Dawn mission to reach more than one target with a significantly smaller launch mass compared to using a chemical propulsion system, which would only be practical to reach a single target [16]. EP can allow a larger payload or a reduced mission budget by reducing the mass lifted to orbit (at a cost on the order of several thousands of dollars per kilogram with current launch systems [17]), or extend the lifetime of satellites by efficient use of propellant for station keeping.

Though the potential benefits of EP systems are significant, there are tradeoffs and limitations. Most EP systems produce thrust under 1 N, much smaller than is possible with chemical rockets. Though high power exceptions with thrust above 1 N do exist (e.g. [18]), all EP systems produce low thrust. The main consequence of low thrust is that maneuvers require a very long burn time, often on the order of weeks or months, to achieve the desired velocity increment.

The main reason for low thrust is that accelerating the propellant to high speed requires a high power input, and the power available from the onboard power supply is limited. The problem only worsens at higher I_{sp} . Since thrust is proportional to exhaust velocity but the power required is proportional to velocity squared, there is a fundamental tradeoff between I_{sp} and the best thrust-to-power ratio that could be achieved by an ideal EP system:

$$\frac{T}{P} < \frac{\dot{m}v_{ex}}{\frac{1}{2}\dot{m}v_{ex}^2} = \frac{2}{v_{ex}}. \quad (2.3)$$

Another tradeoff in that vein is that higher I_{sp} systems will require more powerful and therefore larger and more massive power supplies. Assuming the mass of a power supply is proportional to the output power, a simple analysis [13] shows that the effective propellant mass including propellant and power supply is:

$$m_{p,eff} = m_p + m_{power\ supply} = \frac{T\Delta t}{v_{ex}} + \frac{\alpha T v_{ex}}{2\eta}, \quad (2.4)$$

where T is thrust, Δt is the burn time, α is the constant of proportionality between output power and power supply mass ($m_{power\ supply} = \alpha P$), and η is the efficiency of conversion from power supply power to ion beam kinetic power. There is an optimal I_{sp} that minimizes effective propellant mass because the propellant itself is too massive if I_{sp} is too low and the power supply is too massive if I_{sp} is too high. This effect is known as the power supply penalty. Although EP in general can greatly reduce propellant mass compared with chemical propulsion, a higher I_{sp} is not necessarily better, and mission designers must carefully take into account all variables.

2.3 Elementary Hall Thruster Design and Operating Principles

2.3.1 Magnetic Layer Hall Thrusters

A magnetic layer Hall thruster has a conceptually simple design shown schematically in Figure 2.1. Propellant gas feeds into the discharge channel through the anode at the upstream end of the channel. The discharge voltage, V_d , on the order of hundreds of volts is applied across the anode and cathode. A hollow cathode external to the discharge channel supplies electrons for the discharge. Electromagnets apply an approximately radial magnetic field that reduces electron mobility across the field toward the anode. Reduced electron mobility in the magnetic field causes most of the discharge voltage drop to occur near the region of strong magnetic field to ensure discharge current continuity on average [3]. The insulating channel walls are typically made of boron nitride (BN) for its insulation, heat resistance, and secondary electron emitting properties.

The magnetic field magnitude is set so that electrons are magnetized and drift azimuthally in the fields due to $E \times B$ drift, but magnetic field is sufficiently weak that ions are

unmagnetized. The azimuthal electron drift, known as the Hall current, is the origin of the thruster's name. The Hall parameter is defined theoretically as the ratio of electron cyclotron frequency to collisional frequency [12] or more empirically as the ratio of azimuthal electron current density to axial electron current density [19]. In either case, it is much greater than unity, ensuring a long residence time for electrons. This improves efficiency by increasing the number of ionizing collisions per electron to sustain the discharge and create new ions to be accelerated [19]. Since the relatively massive ions are unmagnetized, they are accelerated out of the discharge channel with a characteristic residence time on the order of microseconds [20]. This time scale implies that collisional effects are not important for ions and the ions can be considered to be in "free fall" in the field. Most of the ion VDF width originates in overlap between the ionization and acceleration regions. The ion beam is neutralized with electrons supplied by the hollow cathode.

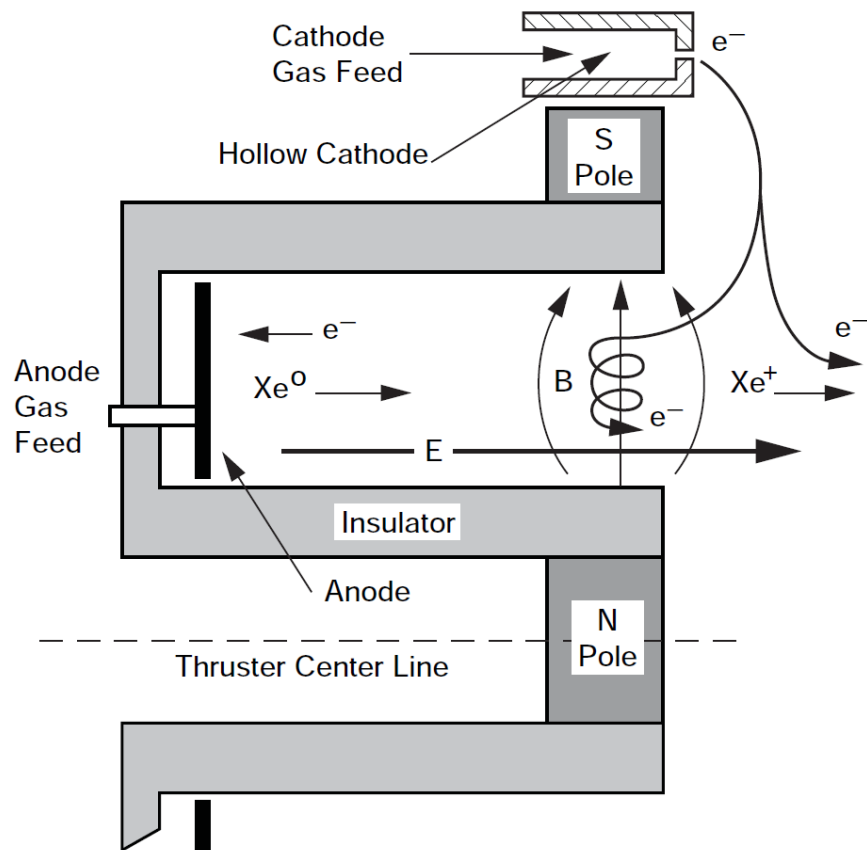


Figure 2.1. Cross-sectional diagram of a basic, magnetic layer Hall thruster. The key parts in construction, electric and magnetic fields, and representative paths of ions, electrons and neutrals are shown. Reproduced from Figure 7-1 in [12].

2.3.2 Orificed Hollow Cathodes

The goal of this section is to describe the cathode's role in a HET and the basic operating principles of hollow cathodes. The discussion is important for this dissertation not only because the cathode is critical to Hall thrusters, but also because the initial validation campaign of the TRLIF system used a hollow cathode as the plasma source.

Virtually all of the power to sustain the discharge comes from the discharge power supply (neglecting the heater and keeper of the cathode and the electromagnets of the thruster), which biases the anode to a few hundred volts relative to the cathode. The (conventional) discharge current from the anode to cathode is necessary to supply the power consumed by the discharge. Ions are mostly born downstream of the anode and those born in the discharge channel are accelerated into the ion beam and do not continue all the way to the cathode. Hence, electrons are critical to carry much of the current to close the circuit. The cathode provides electrons, some of which neutralize the charge and current of the ion beam and some of which contribute current toward the anode. Though the cathode provides some electrons that eventually arrive at the anode, most of the electrons of the anode current are created in ionizing collisions in the region of the Hall current [3].

Early Hall thrusters used a hot filament cathode (e.g. [5]), but modern Hall thrusters use an orificed hollow cathode (commonly shortened to "hollow cathode" or simply "cathode") due to the increased current output and lifetime. In addition, hollow cathodes have lower power requirements since hollow cathode inserts have lower emissive temperatures than hot cathode filaments [12].

The key components of an orificed hollow cathode are shown in Figure 2.2 [12]. It is composed of a hollow cylindrical shell of a thermionically emitting material such as LaB₆ or BaO (the "insert") in a conducting and highly heat resistant tube made of graphite or a refractory metal (the "cathode tube"). There is a gas inlet on the far end of the cathode tube and an orifice typically smaller than the insert radius on the near end. An electrical heating element is wrapped around the cathode tube near the location of the insert.

To operate the cathode, the heater heats the insert to thermionic emission and a potential is applied across the keeper and cathode while gas flows steadily. A DC discharge is lit across the keeper and cathode, with electrons ionizing the neutral gas in the cathode tube, which is at a relatively high density regulated by the orifice diameter. The plasma density in the cathode is

greater than any other part of a Hall thruster, up to the order of 10^{20} m^{-3} [21], [22]. After lighting, the insert is heated by ion bombardment from the very dense plasma inside the cathode tube and the heater may be turned off. Ideally, the cathode is operated so that bombardment by ions accelerated in the sheath sufficiently heats the insert without sputtering the insert material. If there is an external anode in use (such as with a Hall thruster), the keeper power supply may be turned off after the discharge is transferred to the external anode.

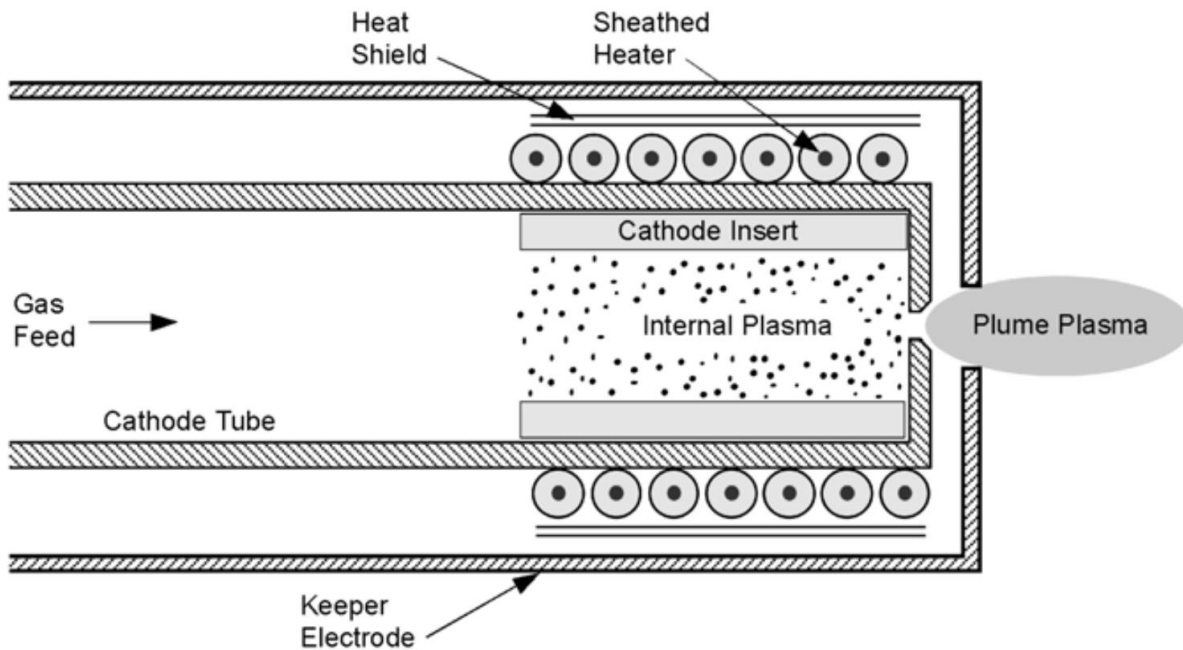


Figure 2.2. Cross-sectional diagram of an orificed hollow cathode. Image reproduced from Figure 6-4 in [12].

Like Hall thrusters, hollow cathodes have been studied and have been in service for many years, but the details of the physical processes at work are not fully understood [21]–[23]. There is also a great deal of current research on how cathodes couple to Hall thrusters and their effects on thruster efficiency. The cathode flow fraction (CFF) is the ratio of cathode mass flow rate to anode mass flow rate. The cathode coupling voltage, the difference of cathode to beam potential, is a useful figure of merit because it indicates the potential necessary for electrons couple to the anode [12], therefore lower values generally indicate better performance. CFF is traditionally 7-10% because early studies showed that high CFF minimizes cathode coupling voltage. In a recent paper [24], Goebel summarized research in this area since about the year 2000 and showed that the total thruster efficiency is in fact constant over a range of CFF values since higher flow leads to lower mass utilization. He argued that lower cathode mass flow may be

optimal since total thruster efficiency is invariant and cathode lifetime is optimized with lower CFF.

2.3.3 Propellant

A wide variety of propellants have been considered [25] and tested on Hall thrusters including mercury, cesium, bismuth [26], iodine [27], and noble gasses. Xenon has become the standard propellant for Hall thrusters and other electric propulsion thrusters for many reasons. It is nontoxic and stray propellant will not condense on or react with spacecraft [12]. It has the lowest first ionization potential of all nonradioactive noble gasses [28]. The heavy atomic mass enables the storage of a higher mass density of gas for a given pressure and gives a greater thrust-to-power ratio than lighter atoms. The lower velocity of heavy ions also implies a higher ion density and electron density by quasineutrality, facilitating sufficient electron transport to the anode to sustain the discharge [3].

Unfortunately, xenon is found in the atmosphere in the trace amount of 87 nl/l [28], making it expensive to extract. For example, I received a quotation in November 2013 of \$22.50 per liter, or almost \$4000/kg, rivaling the cost per mass to launch a payload to orbit. This cost, however, is only significant for laboratory experiments since an actual mission could net a large savings from launching a much smaller weight of propellant compared to the mass necessary for a chemical propulsion system.

2.3.4 Hall Thruster Design Variations

Note that there are many variations to the design of Hall thrusters that optimize different features. For example, a thruster with anode layer (TAL) has electrically conducting channel walls. The reduced secondary electron emission from channel walls results in a higher electron temperature and a shorter acceleration zone, leading to a shorter discharge channel design [1].

A cylindrical Hall thruster (CHT) has a cusped magnetic field geometry with a short annular region followed by a cylindrical region since the inner channel wall is shorter than the outer wall. The design is better for small scale thrusters (channel diameter of a few centimeters or less) in part because it reduces losses to the walls by increasing the ratio of volume to surface area [29] and also because magnetic field scaling is less of an issue [30].

Clusters of identical Hall thrusters can provide a high maximum power, a wide throttling range at high efficiency, and layers of redundancy. One main benefit is also the reduced cost in

research and development compared to a high power monolithic thruster since a cluster uses lower power thrusters that are easier to test with existing vacuum facilities. Clusters can operate with increased thrust than the sum of the individuals and with slightly increased efficiencies due to interactions between the thrusters and their plumes [31].

Nested Hall thrusters have multiple discharge channels of different radii built on the same thruster centerline. Like clusters, they have a high maximum power and large and efficient throttling range due to the different combinations of operation. The X3, a 100-kW class, 3-channel nested Hall thruster developed at PEPL, benefits from a smaller footprint and mass-to-power ratio than a monolithic thruster or clustered configurations of smaller thrusters [18].

An important and relatively recent variation is the development of magnetically shielded thrusters, which have a magnetic field topology designed to reduce interactions sputtering of the walls by reducing both the kinetic and sheath energy and the ion flux to wall surfaces [32]. This design reduces the wall erosion rate by orders of magnitude with minimal efficiency loss [33], thus promising to improve thruster lifetime since the eventual failure by wall erosion has been the main failure mode for Hall thrusters. Unlike typical magnetic layer or anode layer thrusters, ceramic or conducting walls make little difference in the performance of a magnetically shielded thruster [34], providing compelling evidence that wall interactions are not significant for magnetically shielded Hall thrusters.

This dissertation and all the discussion of Hall thrusters hereafter is concerned only with the single channel, unshielded, magnetic layer type of Hall thruster. This is also known as a Hall effect thruster (HET), or stationary plasma thruster (SPT). Future use of the term "Hall thruster" will refer to this type unless otherwise stated.

Although practical Hall thrusters were developed in the 1960s and the first Hall thruster on a satellite was launched by the Soviets in 1971 [2], significant questions still remain about the physics of Hall thrusters. The following sections describe some of the major topics of research in Hall thruster physics that are relevant to this dissertation. Much of current research can be organized in two related categories. One category is to understand the sources of anomalous electron transport, which generally refers to an electron mobility that is observed to be much greater than classically expected. The other is to understand different thruster operating modes, the oscillations and instabilities that occur in each mode, as well as the conditions that bring about a transition from one mode to another.

2.4 H6 Hall Thruster

The H6 is a laboratory model Hall thruster developed jointly by the NASA Jet Propulsion Laboratory (JPL), the Air Force Research Laboratory at Edwards Air Force Base, and the Plasmadynamics and Electric Propulsion Laboratory (PEPL) at the University of Michigan to study Hall thruster physics at multiple institutions on a single, modern Hall thruster design. Each laboratory has its own copy of the thruster, though Michigan and JPL use slightly different nominal operating. It can be operated in the range of 0.6-12 kW for discharge power, 1000-3000 s for specific impulse, and 50-500 mN for thrust [33]. The original H6 achieves a relatively high total efficiency of up to 64% at the nominal 300-V discharge current, 6 kW operating condition and 7% cathode flow fraction and 70% at the 800 V, 6 kW operating condition [33].

It has a centrally mounted hollow cathode with lanthanum hexaboride emitter material. The outer electromagnet comprises eight coils in series uniformly spaced around the discharge channel, while the inner magnet is a single coil. The magnetic field inside the channel has a plasma lens topology, with field lines that are concave and symmetric about the channel centerline, when the ratio of outer magnet current to inner magnet current is 1.12. Magnetic field lines are approximately electrical equipotential lines, and therefore this configuration helps to focus the ions away from the channel walls to reduce losses and beam divergence [35]. The trim coil was unused for this dissertation. With boron nitride channel walls, it is a magnetic layer Hall thruster.

Chapter 7 of this dissertation covers the first test of the PEPL TRLIF technique on the H6. A photograph of the H6 mounted in position for the campaign is shown in Figure 7.2. As a way of simultaneously introducing concepts related to Hall thruster physics and background on the previous work on the H6 Hall thruster, the following Hall thruster physics section focuses largely on results specific to the H6.

2.5 High-Speed Diagnostics

With the mounting evidence that oscillations are important in influencing Hall thruster operation, time-resolved studies have become more commonplace and many high-speed diagnostics have been developed to directly measure oscillations in plasma parameters in Hall thrusters. This subsection is not intended to be a complete survey of all high-speed diagnostics developed for Hall thrusters but only those that are most relevant for this dissertation. The high-

speed dual Langmuir probe system discussed in Subsection 2.5.1 is extremely important to this dissertation since it was the first use of the transfer function averaging scheme with Hall thrusters. High-speed imaging is also covered in Subsubsection 2.5.2 for two reasons. First, because the general results for the spoke instability and how it relates to mode transitions of the H6 motivate the design of the initial campaign of the new time-resolved LIF system with the H6. Second, high-speed imaging analysis was used during that campaign to detect a possible mode transition during a magnetic field sweep. The discussion of subsubsection 2.5.2 is necessary to appreciate those results and understand how the analysis was done.

2.5.1 High-Speed Dual Langmuir Probe

Lobbia and Gallimore [7], [36] developed a high-speed dual Langmuir probe (HDLP) system that allows fully-swept Langmuir probe IV characteristics to be measured on the order of 100 kHz. The system uses specialized filtering and amplifier circuitry and a “null probe” that is insulated from the plasma but is built with nearly identical cabling and position as the true Langmuir probe. The null probe signal can be subtracted from the active probe signal to removed noise and interference from the leakage currents of stray capacitance, plasmadynamic capacitance, and EM pickup, all common impediments to high-speed Langmuir probe measurements [36]. It was found that the high-speed LP system actually has relatively high SNR compared with a traditional LP system that sweeps over the millisecond time scale due to the noise and distortion introduced to the IV trace by rapidly changing plasma parameters [7].

The system was first demonstrated on an operating condition of the BHT-600 with strong breathing oscillations (see Subsection 2.6.2). The plasma density and discharge current oscillated with a peak-to-peak oscillation magnitude of nearly 80% of the mean values with a delay time between the two corresponding to the ion transit time. The traveling wave of ion density was taken as evidence of an approximately toroidal expulsion of ions with each “breath.” The electron temperature and plasma potential oscillated similarly to the discharge current without a time delay, implying a faster, electron-driven process determines them. The oscillation in electron temperature and the absence of a delay from the discharge current also lends some empirical support to the more recent theoretical work of Hara proposing that the breathing mode may be excited by oscillations in electron temperature (discussed in Subsection 2.6.2).

The high-speed Langmuir probe system was a key precursor to the work of this dissertation due to its use of transfer functions to model the plume plasma properties as the

“output” from thruster system resulting from the “input” discharge current signal. At each spatial location, the transfer function itself is measured. The transfer function maps an input signal (e.g. discharge current) to an output signal (e.g. electron density at a point in space). The transfer function analysis enables the synthesis of a characteristic output signal of the linear system that corresponds to a particular input signal. As a part of the signal processing used in this dissertation, the transfer function averaging scheme is described in Subsection 4.3.3 and is not reproduced here.

The main purpose of transfer function averaging for the HDLP system is synchronizing the system outputs at all spatial locations to correspond to a single input signal. With only one probe, the datasets at each individual spatial point were measured serially and therefore each original time series cannot be directly compared or plotted with another in the time domain since oscillations vary in a non-repeatable way. The characteristic output signals as a function of time for all spatial locations are synthesized from the transfer functions using a single input discharge current trace (e.g. the discharge current trace displayed in the line plot of Figure 2.3). It can be shown that the characteristic output signal of the transfer function corresponds to the actual output signal that occurred at the time of the input signal (see Subsection 4.3.3). Therefore, the characteristic output signals generated for a given input signal all correspond to the same time. This allows a coherent visualization of the plume properties across many points in space even if oscillations are non-repeatable and the measurements of the transfer functions at each spatial location were made at different times. Figure 2.3 shows an example of several heat maps of electron density at many different points in space all synchronized in time over a period of oscillation because all characteristic output signals (electron density at each spatial location) were generated for a single input signal (discharge current).

One of the key questions concerning this dissertation is whether there is indeed a linear system relating input signal to output signal, a requirement for the existence of the transfer function. Lobbia provided some evidence to support this assumption by pointing out several features of the output signal that are physically expected or agree with independent measurements. For example, the oscillations in electron density at different axial locations have a high cross-correlation with a time delay corresponding to the ion transit time between the two positions [7], [36]. Also, the ultimate velocity of the ion density wave corresponds within 1% to the thruster I_{sp} measured independently [36].

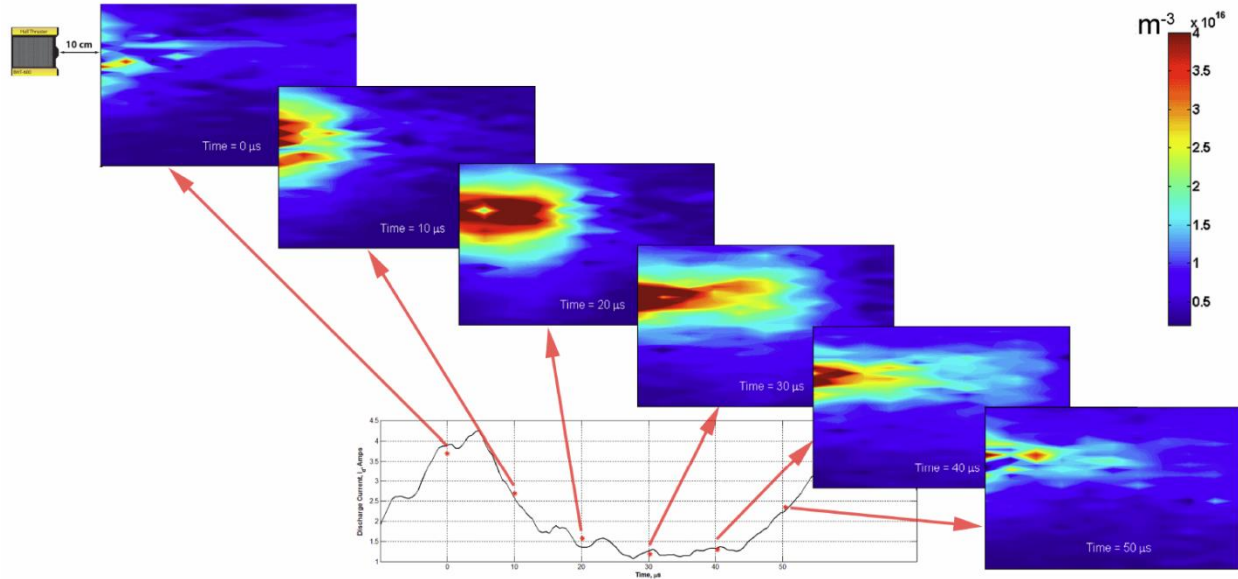


Figure 2.3. Heat maps showing how the electron density changes over one period of a breathing oscillation of the BHT-600. The data are synthesized as the output of the transfer functions at each spatial point for a single input discharge current trace. Reproduced from [36].

The HDLP system has an opportunity to validate the characteristic output signal of the transfer function by comparing directly to the original time series of the output signal from Langmuir probe analysis. Figure 2.4 shows that the characteristic electron density from the average transfer function (red) is nearly identical to the original time series (black) when averaging over 10 or 100 transfer function estimators. Note that the transfer function is measured by averaging over Q different transfer function estimators to reduce noise (see Subsection 4.3.3). The $Q = 1$ case is an exact match by definition since there is no averaging over the transfer function coefficients; they are simply defined to be the ratio of the single output to input discrete Fourier Transform components. The close match of the characteristic output of the transfer function to the original signal unambiguously validates the transfer function approach.

As mentioned previously, the direct comparison between the characteristic signal and the original signal that it models is possible for the HDLP system due to the high signal-to-noise ratio of the raw time series following LP analysis. It is not possible for LIF results because transfer function averaging is the primary averaging scheme to recover the signal from the background noise, and the raw data are completely buried in noise before transfer function averaging. **It is for this reason that much of the dissertation is concerned with various**

benchmarks to help validate the accuracy of the characteristic LIF signal from the transfer function.

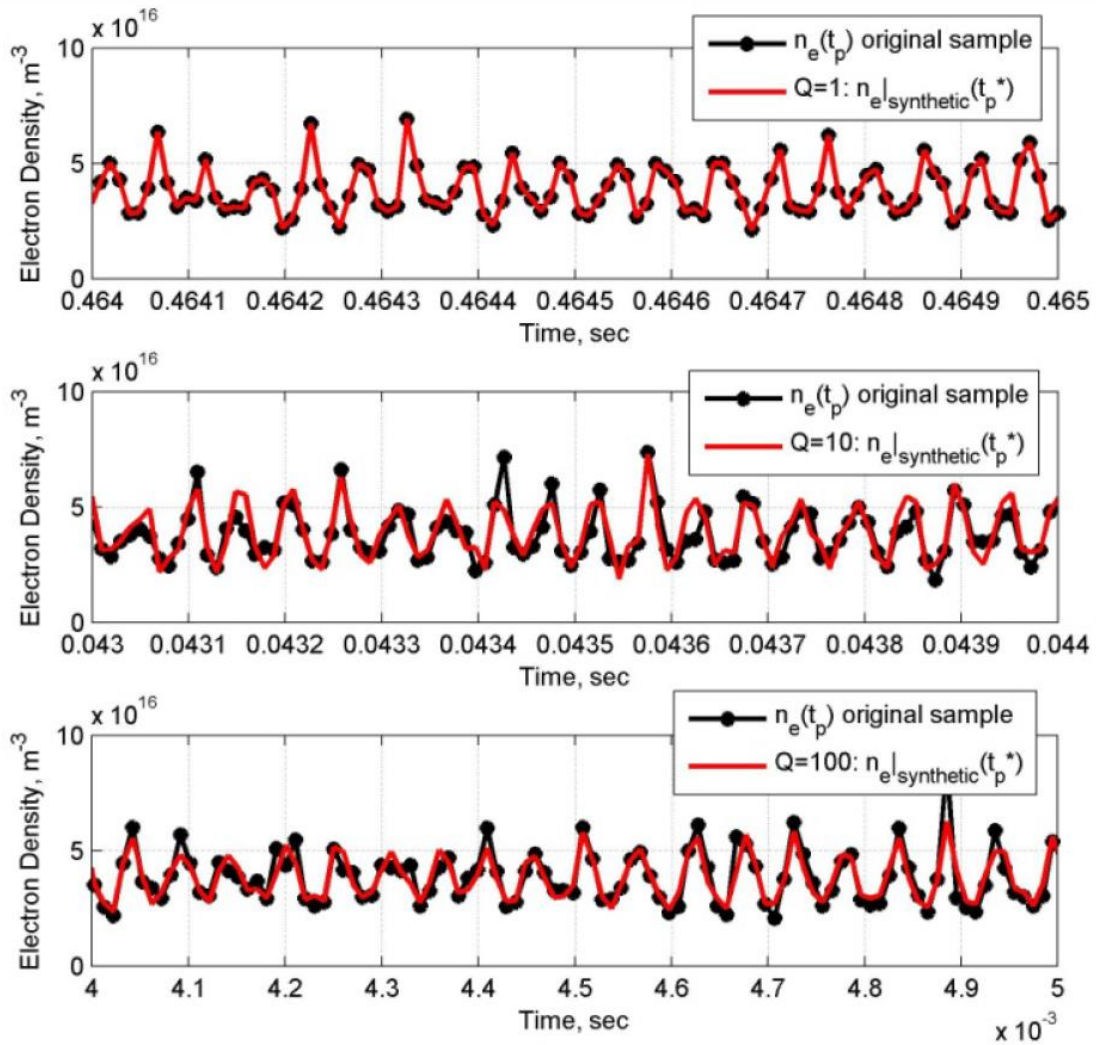


Figure 2.4. The synthesized characteristic electron density from the average transfer function (red) reproduces the key features of the original signal (black). Reproduced from [37].

2.5.2 High-Speed Imaging and the Spoke Instability

Capturing high-speed video of the light emission from a Hall thruster has emerged as a powerful, though qualitative, diagnostic of thruster operation. The main advantage of the use of a high-speed camera is the ease with which it can collect information about events occurring simultaneously at all azimuthal and radial positions. This contrasts with most other probe and optical diagnostics that typically measure plasma properties within a small volume or a single line-integrated path, and are therefore typically give an incomplete picture of the plasma

properties in the discharge or plume. Therefore, the high-speed camera is most useful to characterize azimuthal instabilities, primarily the low-frequency spoke instability.

Though others have also investigated thrusters with high-speed imaging (e.g. [38], [39]), the discussion here focuses on the work of McDonald, and Sekerak. A portion of this dissertation in Chapter 7 involving high-speed imaging is related to their work by both the H6 Hall thruster and the high-speed imaging analysis scheme, therefore their work provides the most relevant background.

The spoke instability refers to one or more regions with brighter than average light emission rotating azimuthally around the discharge channel in the $E \times B$ direction. Spokes are generally within the radial extent of the discharge channel, but can be observed between the channel and a centrally mounted cathode [8]. Azimuthal instabilities in Hall thrusters have been known and studied to a certain degree for decades [4], but the application of high-speed cameras is relatively new. Spokes have been observed across a wide range of Hall thrusters [8] and vary widely in terms of amplitude, frequency, coherence, and the number of spokes depending on thruster operating conditions [8], [40]. The onset of strong spokes in the H6 was found to be related to a transition from a high-current/global thruster mode to the low current/local mode as discussed in Subsubsection 2.6.1.2.

Spokes are suspected to contribute to anomalous electron transport since azimuthal perturbations in plasma properties could lead to azimuthal electric field, which would generate an axial $E \times B$ drift. McDonald used the approach of Janes and Lowder [5] to estimate that the spokes carry nearly half of the total current [6]. If true, that would explain much of the anomalous electron transport at some thruster operating conditions. Note that a time-resolved LIF system could potentially detect azimuthal electric field perturbations and provide another method to verify calculations such as these.

The high-speed image analysis technique was described in detail by McDonald [8] and later modified by Sekerak [10]. The basic analysis scheme can be succinctly described as follows. The raw video is often taken at a resolution of 256×256 at a frame rate of 87.5 kHz for a short length of time up to about 1 s at most. The resulting data can be considered a 3D array of dimensions $256 \times 256 \times N_f$, where N_f is the number of frames. The signal is AC coupled by subtracting the pixel-wise mean image from each frame. The discharge channel annulus is divided into 180 azimuthal bins of 2-degree extent. For each frame of the video, the average

intensity of all pixels in each bin is taken to be the bin intensity. The resulting 1D array of length 180 representing bin intensity as a function of the azimuthal angle for each frame in time is illustrated in Figure 2.5. This process converts the 3D array representing the whole video into a 2D array of dimensions $180 \times N_f$.

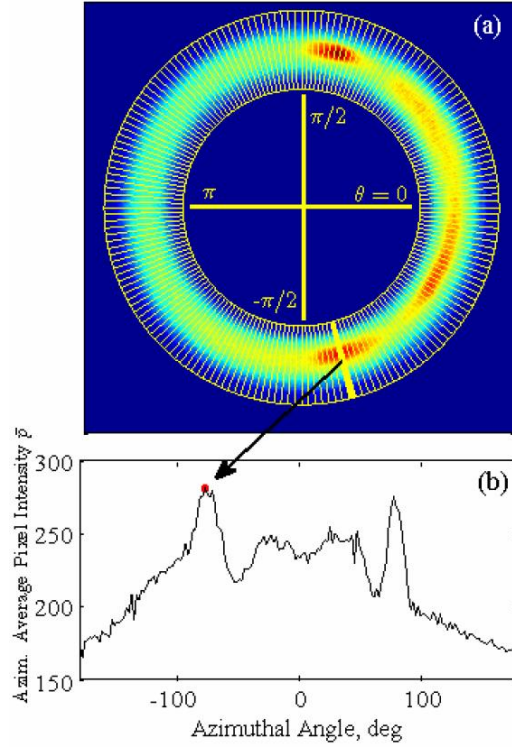


Figure 2.5. The 2D array of a single frame of the video is azimuthally binned and the intensity is averaged over 180 bins to generate a 1D array of bin intensity as a function of azimuthal angle. The complete video is then represented by a 2D array of dimensions $180 \times N_f$. Reproduced from Ref. [40].

A 2D discrete Fourier transform (DFT) gives Fourier coefficients for a basis set of sinusoidal functions with frequencies and wavenumbers up to the Nyquist frequency and Nyquist wavenumber. The top frame of Figure 2.6 shows an example of the 2D power spectral density, another 2D array of the same dimensions. It is proportional to the magnitude of the DFT coefficients (which are complex numbers). The wavenumbers corresponding to an integer number of periods around the thruster annulus are of particular interest and denoted by the mode number m . The mode number corresponds to the number of spokes rotating around the discharge channel. The $m = 0$ mode is defined to correspond to zero wavenumber. It corresponds to oscillations as function of time that do not depend on space, or in other words it captures fluctuations of light intensity over the whole channel. In general, there is a mix of many different

modes occurring simultaneously. A 1D array for each relevant spoke mode can be taken out of the 2D array of power spectral density. This last simplification allows the power spectral densities of the spoke modes to be plotted in line plots for a simple and intuitive comparison with each other and the power spectral density of the discharge current, as illustrated in the bottom frame of Figure 2.6.

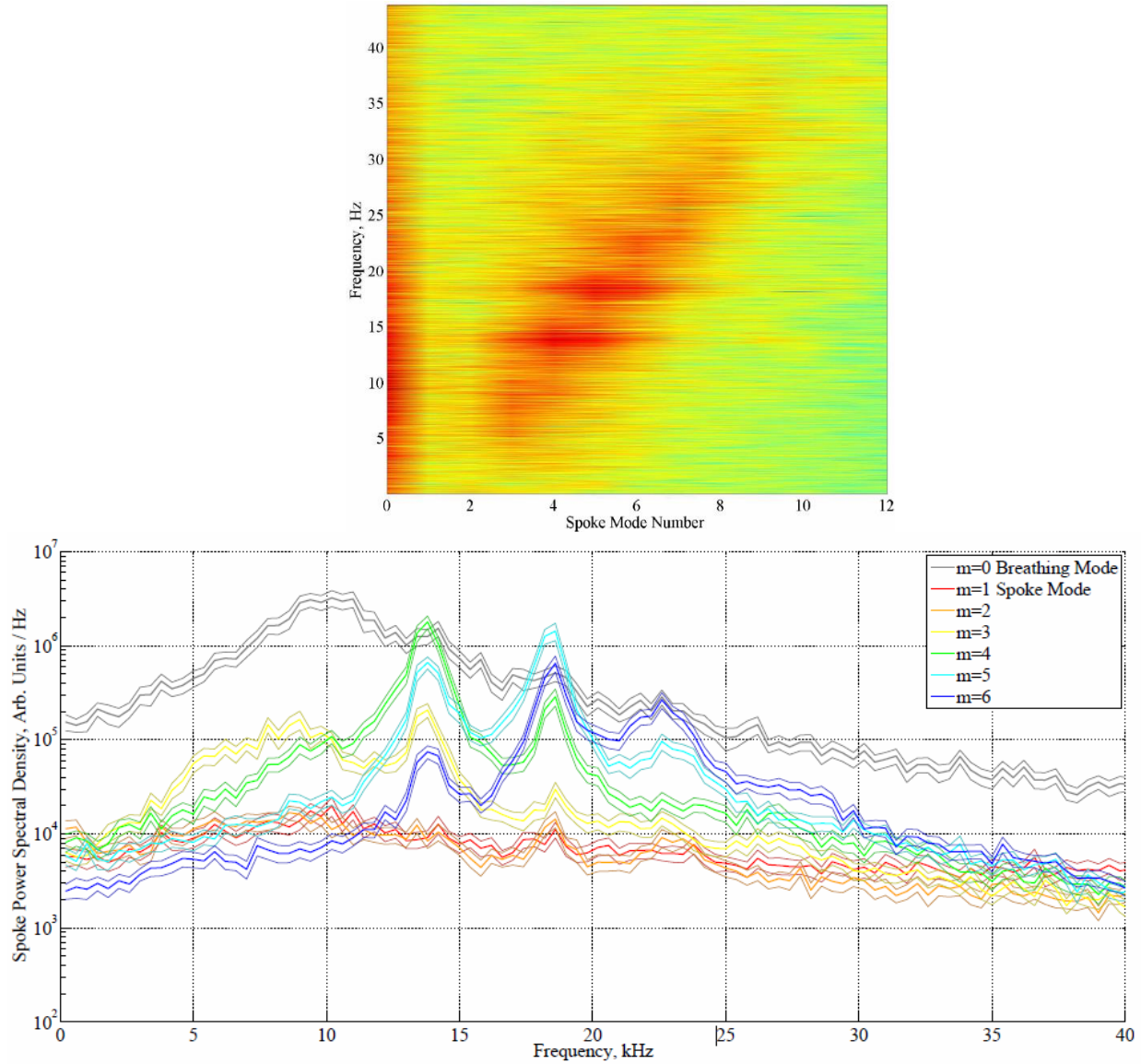


Figure 2.6. Top: example 2D power spectral density of a $180 \times N_f$ array of azimuthal intensity as a function of time. Bottom: vertical cross-sections of the 2D power spectral density are taken for a number of spoke modes with $m > 0$ and the $m = 0$ mode. Reproduced from Figure 6 of Ref. [6].

High-speed imaging has yielded several interesting trends for the spokes observed. An increase in magnetic field strength tends to make spokes stronger and more stable, with higher

spoke modes appearing and becoming dominant and their velocity decreasing [8], [9], [40]. Higher spoke modes travel faster than lower modes. The spoke strength relative to breathing mode reaches a maximum with operating conditions near settings corresponding to optimal thruster performance with the H6 [40], though that is not always true for other thrusters.

The light intensity in the high-speed video has been shown to be intimately related to the discharge current. High-speed video taken with the H6 modified with an azimuthally segmented anode verified that local current to a segment is strongly correlated to the spokes, with all of the same peaks appearing in the $m > 0$ power spectra [6]. The variation of current at the anode segments attributable to spokes actually dominated that of the breathing mode, but power spectral density of the total discharge current of all the segments or the discharge current of an unsegmented anode show little or no peaks at spoke mode frequencies since azimuthal variation tends to be averaged out. In addition, the power spectral density of the average optical emission of each video frame [37] and the $m = 0$ spoke mode [40] both correlate very well with the discharge current power spectrum. The RMS pixel fluctuation of the camera also correlates very linearly with RMS discharge current over many operating conditions (see Figure 10 in [40]), unsurprisingly since the RMS can be found by integrating over the power spectral density.

One drawback of the high-speed image analysis technique is that there is no precise physical interpretation of the data corresponding one-to-one to a single plasma property. The light emission is usually unfiltered and corresponds primarily to emission from electron-impact excitation, but its intensity can vary according to electron energy distribution and densities. Estimates of time-resolved electron temperature or other plasma properties might be possible by filtering light by wavelength, capturing multiple lines with the high-speed camera, and applying a collisional-radiative (or simpler) model. This approach would essentially use the high-speed camera to quickly perform a measurement of optical emission spectroscopy with spatial and temporal resolution. Dale and Gallimore [41] recently reported initial tests in this vein, but noise prevented useful time-resolved estimates of electron temperature. Based largely on the relationship of the high-speed video to the discharge current discussed above, Sekerak argued that the instantaneous light emission from a region of the discharge channel is likely strongly correlated with the discharge current density in that region. To help bridge the gap to a measurement of a physical property, he therefore scaled the light intensity to units of discharge current density [10], but the accuracy of that estimate remains unclear.

2.6 Hall Thruster Physics

2.6.1 Mode Transitions

2.6.1.1 Definition of Hall Thruster Mode Transitions

A hall thruster operating condition is defined by the set of mass flow rate, discharge voltage, and magnetic field topography and magnitude. Some thruster designs allow some adjustment of the field topology due to the presence of multiple electromagnets, but authors often hold some nominal field topology constant while varying the magnitude alone by changing current in all electromagnets by the same factor. This definition of an operating condition is a useful concept because the thruster's performance (e.g. thrust, efficiency, I_{sp}) and plasma dynamics (e.g. plasma properties, oscillation spectra, and instabilities) are typically found to be a function of these parameters [42], [10]. Nevertheless, hysteresis is sometimes observed, for example, where the thruster can operate in either a high or low discharge current regime at a single operating condition by either raising mass flow rate and magnetic field magnitude from lower values or by decreasing them from higher values, respectively [43].

There is no precise definition of Hall thruster operating modes (sometimes called regimes) or of mode transitions. The general idea common among authors is that changes to operating conditions yield gradual changes in measured properties and typically follow a consistent trend within the bounds of a single mode. Authors use the term "mode transition" to refer a sharp change in measured properties, such as mean discharge current or oscillation spectra, that occurs over a small change in operating condition parameters. Another key feature of a mode transition is a departure from some preceding trends that accompanied changes in operating condition parameters.

In 1976, Tilinin published one of the most complete studies of Hall thruster mode transitions. A laboratory model "Hall accelerator" that predates modern Hall thruster designs was tested, leading Sekerak to point out that the results may have limited applicability to current thrusters [44]. The results are nonetheless instructive as a complete study of mode transitions and there are distinct parallels to mode transitions found in modern Hall thrusters. Tilinin found that measured oscillations were only weakly dependent on discharge voltage and mass flow rate and carried out a study of oscillatory behavior as a function of magnetic field magnitude. High-frequency oscillation amplitude (on the order of megahertz and higher) was found be related to

low frequency (kilohertz) amplitude oscillations such that an abrupt change in one often accompanied a change in the other, though the correlation is not always positive or negative.

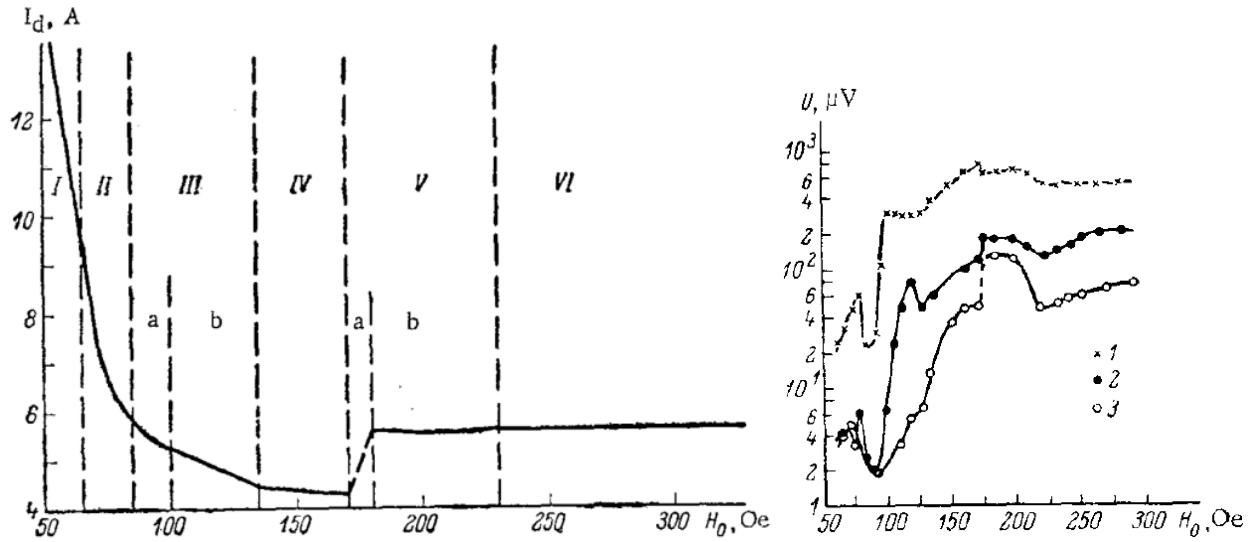


Figure 2.7. Left: Plot of mean discharge current versus magnetic field magnitude over the six mode regions. Right: High-frequency oscillation amplitude plotted as a function of magnetic field magnitude for frequencies of (1) 20; (2) 60; (3) 140 MHz at a floating probes just downstream of the exit plane. Reproduced from figures 5 and 4d in [42].

Six distinct modes were identified and shown in the plot of discharge current versus field magnitude on the left side of Figure 2.7. In the "collisional (classical) conductivity" regime (I), the cycloid path of electrons is comparable to the dimensions of the device and classical collisional conductance explains the observed conductance. High and low frequency oscillations are very low amplitude. An azimuthally propagating wave at approximately the $E \times B$ velocity appears in the "regular electron drift wave" regime (II). Low conductivity causes loop oscillations to appear in the "transitional" regime (III). The "optimal regime" (IV) is so called because the ratio of ion beam current to discharge current is maximized here. Low frequency oscillations are minimized. Regime V is called the "macroscopic instability" regime because there are oscillations of low enough frequency to be visible by eye and strong enough to spontaneously extinguish the discharge. There is a distinct jump in discharge current and the electron drift wave is not detected. The thruster becomes stable again in the "magnetic saturation" regime (VI), where high frequency oscillations tend to approach a maximum amplitude and low frequency oscillations are minimal.

2.6.1.2 Mode Transitions of the H6 Hall Thruster

Brown and Gallimore [43] discovered transitions between two distinct operating modes on the H6 Hall thruster (see Section 2.4) in a study of operation at low discharge voltages of 100-120 V and anode mass flow rates of 10-20 mg/s. The thruster transitioned from a "low-current" mode to an undesirable "high-current" mode as discharge voltage was reduced. A lower CFF lead to a transition to high-current mode at a higher voltage, but auxiliary neutral flow outside of the cathode could stave off the transition.

The two modes were called "low-current" and "high-current" modes due to an approximately 10% difference in mean discharge current. The high current mode also featured larger amplitude discharge current oscillations by about a factor of ten. The other key features that Brown and Gallimore [43] found to distinguish them are that the high-current mode has generally poorer performance in terms of a lower thrust-to-power ratio, lower current utilization (ratio of ion beam current to discharge current), and a larger beam divergence. A hysteresis effect enabled operation in both modes at some operating conditions. Though the amplitude of oscillations changed in different modes, the peak frequencies were the same in both modes at the same operating condition. Brown also observed a sudden visual change in the plume at the transition, shown in Figure 2.8. The luminous plume in high-current mode extends further downstream, has a more prominent central spike, and is generally brighter.

Using an $E \times B$ probe and a retarding potential analyzer, measurements of the ion energy and ion energy per charge distributions in the far field plume were well defined in low-current mode but highly diffuse in high-current mode, possibly even showing bimodal distributions. Though the ion distributions clearly differed between the two modes, gaining quantitative information about ion distributions was difficult since the exact contributions from different ionization states were unknown.

LIF would help uncover more information about the ion distributions in the two modes because it can be spatially resolved in the discharge and near-field plume (e.g. [45], [46]), and the ion velocity distribution is measured directly without confounding multiple charge states. Bimodal ion velocity distributions have been observed at a number of operating conditions of the H6 (e.g. [45]). A new TRLIF technique would enable measurements of oscillations in the distribution, possibly shedding light on the apparently bimodal form of the ion energy and velocity distributions.

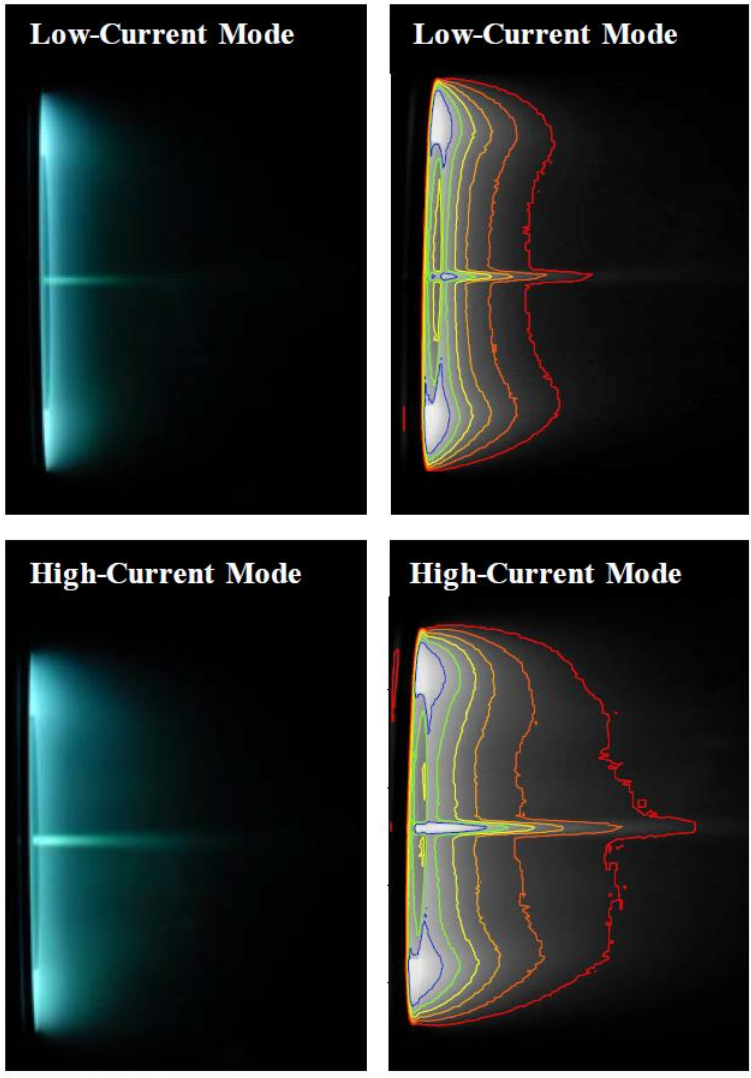


Figure 2.8. Photos of the H6 plume (left) and contours of constant intensity (right) for both low-current mode (upper) and high-current mode (lower) [43]. Reproduced from Figure 1.

Sekerak et al. [10] conducted a study of thruster mode transitions as a function of magnetic field magnitude at discharge voltage of 300-450 V and anode mass flow rate of 14.7-25.2 mg/s. Sekerak found transitions between two modes with the same general features as the modes observed by Brown for low voltage conditions. The higher current mode (by about 15% in mean discharge current) was again found to be related to ten times higher amplitude discharge current oscillations (almost 100% of the mean value) and a 25% lower thrust-to-power ratio. Figure 2.9 shows the trends in discharge current and oscillation amplitude as a function of magnetic field. Very similar changes in light emission in the plume were also observed. The higher current mode has generally brighter emission, a more prominent central spike, and more

emission that occurs further downstream. In contrast with Brown's findings, the thruster was more sensitive to a transition to a high-current mode for higher discharge voltage and higher anode flow rate operating conditions [10].

Sekerak et al. [10] also conducted an extensive test of oscillations in the plume during the two modes using a high-speed camera and Langmuir probes. In general, the higher-current mode was found to be associated with a global oscillation in phase at all azimuthal locations, and the lower-current mode was associated with localized, azimuthally propagating spokes (see Subsection 2.5.2), which lead to the terms "global mode" for the higher-current mode and "local mode" for the lower-current mode. The global oscillations are the breathing oscillations described in Subsections 2.5.1 and 2.6.2.

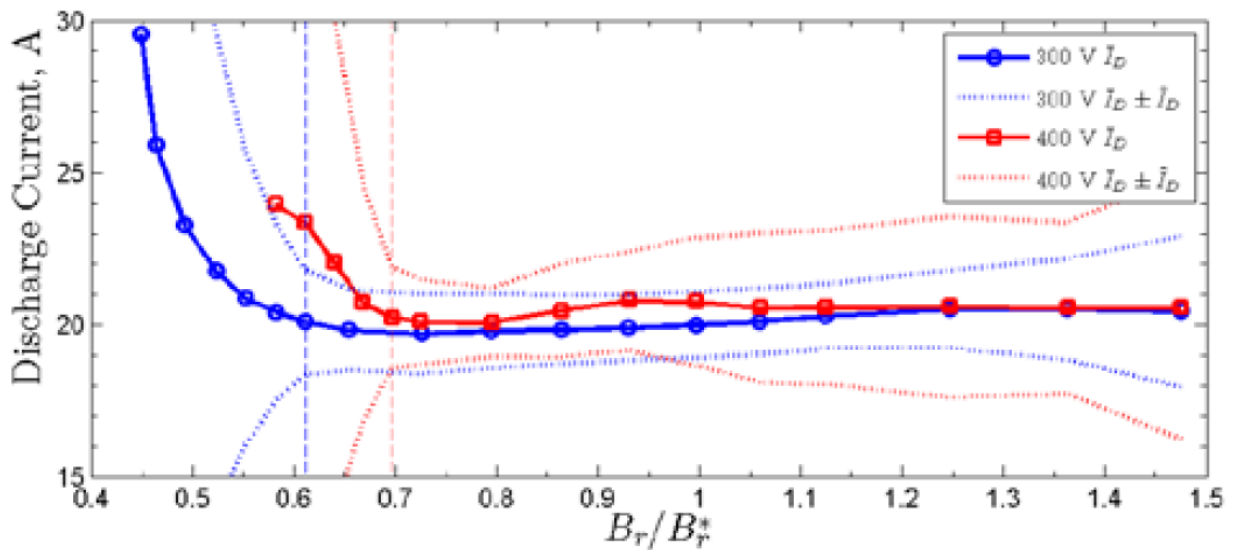


Figure 2.9. Examples of magnetic field sweeps showing transitions in mean discharge current (solid lines) and RMS oscillation amplitude (dotted lines) as a function of magnetic field magnitude from Sekerak et al. [10]. Reproduced from Figure 20 (a).

Sekerak [44] found trends in three quantities to be sufficiently repeatable at all operating conditions tested to define guidelines for the mode transition in terms of certain criteria. One, the ratio of the $m = 0$ peak (PSD_{m_0}) to the strongest $m > 0$ peak in the high speed image analysis ($PSD_{m_{max}}$) (see Subsection 2.5.2 for further discussion of spoke modes). Two, the RMS discharge current oscillation amplitude (\tilde{I}_D) normalized by the mean discharge current (\bar{I}_D). Three, the normalized difference of the average discharge current at the given operating condition (\bar{I}_D) from the minimum average discharge current during the magnetic field sweep

($\bar{I}_{D_{min}}$). Table 2.1 summarizes the conditions determined by Sekerak. In short, the thruster is in "global mode" when: (1) the $m = 0$ peak is at least fifty times larger than all other spoke mode peaks in the FASTCAM power spectrum, (2) the discharge current RMS oscillation amplitude is at least 15% of the mean value, and (3) the mean discharge current is more than 2% greater than the minimum mean discharge current during the magnetic field sweep. Conversely, the thruster is in local mode when: (1) the $m = 0$ peak is less than 10 times the largest spoke mode peak, (2) the RMS current oscillation amplitude is less than 10% of the mean, and (3) the mean discharge current is less than 2% larger than the minimum mean current. Otherwise, the thruster is in a transition region and possibly exhibiting the characteristics of both modes simultaneously or alternating between them, and therefore a specific transition point was difficult to define.

Table 2.1. A summary of the quantitative criteria for defining H6 mode transitions in terms of FASTCAM power spectrum peaks and discharge current magnitude and oscillation amplitude developed by Sekerak [44]. Reproduced from Table 4.2.

Criteria	Global Mode	Local Mode
$\frac{PSD_{m_0}}{PSD_{m_{max}}}$	$\gtrsim 50$	$\lesssim 10$
$\frac{\tilde{I}_D}{I_D}$	$\gtrsim 0.15$	$\lesssim 0.10$
$I_{D_m} = \frac{\bar{I}_D - \bar{I}_{D_{min}}}{I_{D_{min}}}$	$\gtrsim 0.02$	$\lesssim 0.02$

Unlike Brown and Gallimore [43], Sekerak et al. [10] observed an absence of hysteresis when inducing mode transitions by changing operating conditions, meaning the thruster mode was a function of operating condition parameters and did not depend on the path taken in changing parameters to arrive at the final operating condition. Note that Sekerak induced mode transitions by changing magnetic field magnitude while Brown induced them by changing anode flow and discharge voltage. It is unknown whether the difference in hysteresis is related to the difference in operating conditions tested, the different methods of inducing mode transitions, or possibly some other difference between the two campaigns.

This subsection has introduced the two main operating modes of the H6 and briefly described the oscillations that occur in the two modes. Though the breathing oscillations of the global or high-current mode are reasonably well understood (see Subsection 2.6.2), the rotating spoke oscillation and the processes that lead to a transition from one mode to the other are not as

well understood. This deficit and the significantly reduced performance of the global mode motivate the development of time-resolved diagnostics to measure the processes at work in the two modes and during transitions.

It is of particular interest to understand the modes and transitions better because, as Sekerak [10] notes, the transition tends to occur near the setting with an optimal thrust-to-power ratio, and small changes over time to the magnetic field or the discharge channel wall could potentially change the thruster operating mode. In addition, considering that Brown and Gallimore [43] found that the location of the transition changes with auxiliary gas flow, it is possible that facility effects could change the operating mode between ground testing and actual operation in flight. The initial campaign of the new time-resolved LIF system on the H6 in Chapter 7 was designed similarly to the campaign of Sekerak et al [10]. It included a magnetic field sweep and measurements of time-resolved LIF, high-speed video, and discharge current oscillations to characterize the change in ion dynamics during a thruster mode transition.

2.6.1.3 Similar Mode Transitions in Different Hall Thruster Designs

Similar transitions and modes have been described by a number of authors for a variety of Hall thruster designs. For example, an ATON Hall thruster, having completely different neutral gas distribution and magnetic field shape with a null point upstream of the exit plane [47], was found to have two very similar modes [48]. The mode corresponding to "low-current" mode was called "swallow tail" mode due to the prominence of two points near the edges of the plume, and "high-current" mode was called "spike mode" due to the prominence of the spike along the thruster centerline; both features are also easily visible in the H6 in Figure 2.8. Many other features that parallel the findings of Brown and Sekerak in the H6 were reported, such as a 69% larger divergence angle in spike (high-current) mode, and light emission from deeper within the channel in swallow tail (low-current) mode. They concluded that the visual difference was largely due to the interaction of intersecting ion jets from different parts of the channel, which occurs further downstream for swallow tail mode, where there is less light emission.

Yamamoto found a "thruster with anode layer" (TAL) (see Subsection 2.3.4) to have distinct modes and transitions [49]. Similar to Sekerak's findings, "regime 1" at relatively low magnetic field strength featured high amplitude discharge current oscillations, while "regime 2" at slightly higher magnetic field magnitude featured a slightly smaller mean discharge current and greatly reduced oscillation amplitude. Regime 2 was also similar to low-current mode with a

maximum in thrust efficiency and minima in plume divergence and wall erosion rate. Yamamoto also explored "regime 3" at higher magnetic field, where the mean discharge reaches a constant value as a function of magnetic field, while discharge current oscillations again reach large amplitude and discharge extinguishes at a critical magnetic field value, similar to the high magnetic field regimes described by Tilinin (see 2.6.1.1).

These examples demonstrate that qualitatively similar operating modes and transitions are found across a variety of Hall thruster designs. Though the details of the designs surely affect many quantitative differences, this is strong evidence that the physical phenomena associated with these modes and transitions are common to Hall thrusters in general, regardless of the details of the thruster design.

2.6.2 Breathing Mode Theory and Simulation

A Hall thruster discharge often exhibits low-frequency oscillations in the band of 10-30 kHz. These oscillations are historically called "loop" or "circuit" oscillations in Russian literature because they are highly sensitive to the properties of the power supply and can be damped by appropriate power supply filtering [4], [50]. Though known since at least the 1970s, significant research into the physical mechanism involved has only been carried out since the 1990s [50]. Broadly speaking, the breathing mode involves the discharge current and plasma density oscillating in phase and an out-of-phase oscillation in neutral density.

There are two competing interpretations as to the mechanism that sets the frequency. One is an ionization instability where the oscillation is generally considered to be a standing wave and the frequency is set largely by the ionization rate and densities alone (independent of particle transport), and the other is a convective interpretation that the frequency is determined by the time neutrals take to refill the discharge channel after an ionization event depletes the neutral density.

The simplest and most well-known ionization instability theory of the breathing mode dates to the predator-prey model of Fife et al. [51] in the late 1990s. The model takes a simple 0D perspective (it considers the ionization zone as a box of length L with plasma parameters varying only in time and not space). Ions are all created within the volume by ionization and leave at the nominal ion velocity V_i . neutrals enter the zone at the nominal neutral speed V_n and ionization is the only loss mechanism. The conservation equations for ion density n_i and neutral density n_n under these conditions are:

$$\begin{aligned}\frac{\partial n_i}{\partial t} &= \xi(T_e)n_i n_n - n_i \frac{V_i}{L} \\ \frac{\partial n_n}{\partial t} &= -\xi(T_e)n_i n_n + n_n \frac{V_n}{L},\end{aligned}\tag{2.5}$$

where $\xi(T_e)$ is the rate constant for electron impact ionization, which is a function of electron temperature. These equations can be linearized and combined into the form of an undamped harmonic oscillator with the frequency:

$$\omega = 2\pi f = \sqrt{\xi(T_e)^2 n_{i,0} n_{n,0}} = \sqrt{\dot{n}_{i,0}^2 / n_{i,0} n_{n,0}},\tag{2.6}$$

where $\dot{n}_{i,0}$, $n_{i,0}$, and $n_{n,0}$ are the equilibrium values for ionization rate, ion density, and neutral density, respectively.

The model provides testable predictions useful for comparing with both numerical models and experiments. Fife compared the model's prediction with one of the first numerical models to use a physical model for enhanced electron mobility via wall interactions instead of a nonphysical tunable parameter for anomalous mobility. The predicted frequency was close to the result of the numerical model, but there was a significant discrepancy from experimental results that was thought to be due to the numerical model under-predicting electron temperature.

Boeuf and Garrigues [52] simulated low-frequency oscillations for the SPT-100 by treating ions with a collisionless Vlasov equation and electrons as a fluid. The most physical results were found by adding a conductivity term from electron-wall collisions as well as a corresponding electron energy loss term from wall collisions. Without conductivity from electron-wall collisions, the classical conductivity alone is not large enough where the neutral density is low. Furthermore, without an energy loss term from wall collisions, the ionization rate is always too high, leading to ionization deep inside the channel. Bohm conductivity was also tried but was not able to match experimental results as well as the model with electron-wall interactions.

The electron-wall collision model was not completely from first principles and included a tunable parameter, to which the oscillations were highly sensitive. Even so, the sensitivity of the results to electron-wall interactions and the qualitative agreement with experimental results reinforce the notion that Hall thruster operation is strongly influenced by electron-wall interactions. Further support for that general conclusion is that Gascon et al. [53] observed

changes to thruster operation depending on wall material. In addition, Hofer et al. [33] found that magnetic shielding, which greatly reduces plasma-wall interactions, significantly changes the power spectral density of the discharge current and increases oscillation amplitude by 25%, among other changes to thruster operation. The Boeuf and Garrigues model was tested extensively under a variety of conditions. Changes in magnetic field magnitude and discharge voltage influenced oscillations. In general, either a stable regime with no oscillations or an unstable regime with low-frequency oscillations near 15-20 kHz were found, depending on the combination of magnetic field magnitude and discharge voltage. The frequency was an increasing function of discharge voltage but a decreasing function of magnetic field. It was particularly sensitive to the magnetic field profile. A field profile with a small tail of less than 10% that extends nearly to the anode allowed oscillations of a different character near 80 kHz.

At the operating conditions where the 15-20 kHz oscillations were found, the model predicted a periodic depletion and replenishment of neutrals near the exit plane. The high electron temperature and ionization rate quickly ionizes the neutral flow in this region, and the front of neutral flow recedes upstream. The neutral depletion causes the ionization rate and ion density to fall. With a low ionization rate, the neutral front is then allowed to return near the exit plane, which makes another large ionization burst possible. This cycle was described as "breathing," apparently the first use of the term for low-frequency Hall thruster oscillations. This explanation is similar to the predator-prey interpretation, but a key difference is that Boeuf and Garrigues emphasized the role of neutral flow, suggesting the oscillation frequency should scale neutral flow speed, or gas temperature. The numerical simulations also predicted increasing frequency with gas speed. This contrasts with the Fife model that neglects flow entirely with a 0D model. The most relevant results reported by Boeuf and Garrigues that could be testable by time-resolved LIF would be the standing wave of ion density and a slight oscillation in the location of the electric field, which are reproduced in Figure 2.10.

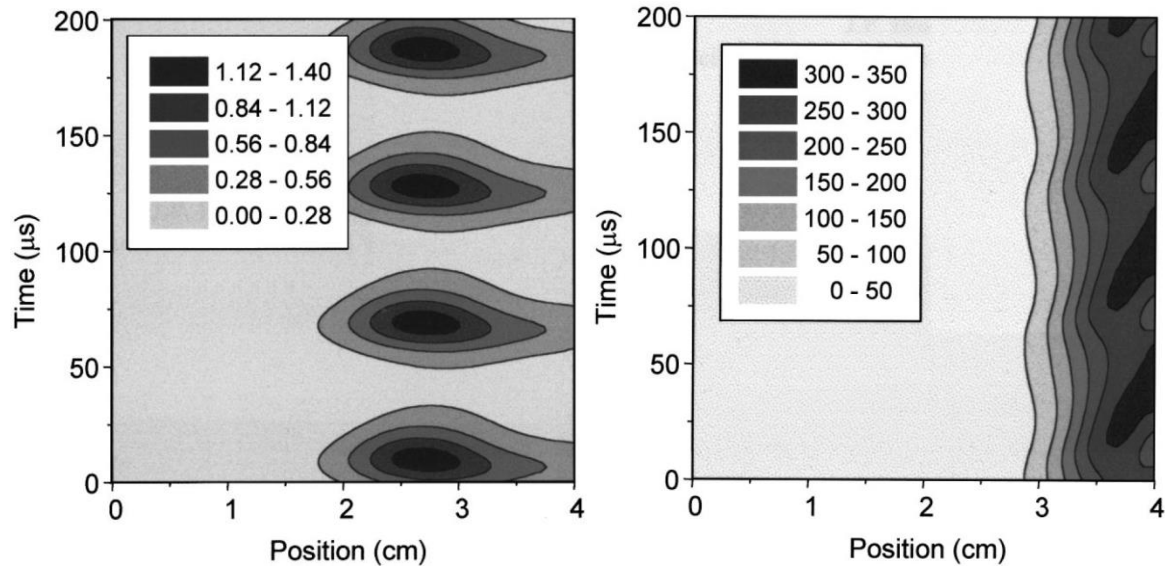


Figure 2.10. Examples of the standing wave in ion density (left, in units of 10^{12} cm^{-3}) and slight oscillation in the position of the electric field (right in units of V) from the numerical model of Boeuf and Garrigues [52]. Reproduced from Figures 9 and 10.

Barral and Ahedo [50] developed a more rigorous analytical model around the mid-2000s. The mathematically complicated details of the theory need not be reproduced here¹, but it is a self-consistent model starting from 1D fluid equations and does not use imprecise parameters, such as the ionization zone length L . To work toward an analytical prediction of frequency, a number of successive approximations are applied. The fluid equations are approximated by a "quasi-static" model in which terms relating to change of the ion distribution are assumed negligible, and the QS model is then linearized. A further approximation for short wavelength oscillations was shown to reproduce the frequency from a numerical simulation.

The main advantage of the Barral model is that it starts from first principles and does not assume either the standing wave or traveling wave interpretations. The theory bares out that there are both standing wave and traveling wave components in the solution, but only the standing wave determines the oscillation frequency to first order [50], [54], contradicting the convective interpretation. To first order, Barral et al. [54] showed that the frequency has the same scaling as the Fife model, leading to an interpretation that the Barral model is a more complete version of the simple predator-prey model. A shortcoming of the theory is that the analytical form of the frequency depends on a "discharge current sensitivity function" that has no analytical form and is

¹ And the author will not feign a complete understanding of the complicated mathematical details.

calculated in the model, but is unknown for a real thruster, hence the theory cannot be directly compared with experiments.

Hara et al. [55] investigated how the electron transport and energy balance are affected by the mode transition between breathing oscillations and a stable mode using a 1D direct kinetic model treating ions and a fluid model for electrons, an approach known as a hybrid-direct kinetic simulation. Firstly, the ionization cost, or the ratio of energy lost to all processes compared to the energy lost by ionization, is somewhat reduced near the anode in the breathing mode. In other words, ionization is suppressed near the anode in the stable mode. Secondly, the energy balance changes. Increasing magnetic field magnitude tends to reduce the axial electron drift velocity, reducing the electron energy loss rate from convection near the exit plane. The reduced kinetic energy from axial and azimuthal drift contributes to a higher thermal energy in the stable mode. The higher temperature leads to the formation of a space charge saturated sheath and causes much larger energy loss to the wall than with the normal sheath that occurs in the breathing mode, and wall loss replaces convection as the main balance to Joule heating. These changes are visualized in Figure 2.11. These findings that the mode transition is associated with changes in electron transport and energy balance suggest that the stability or excitation of the breathing mode may actually be determined by the dynamics of the electrons, despite the fact that the breathing mode is ostensibly an instability between the heavy species.

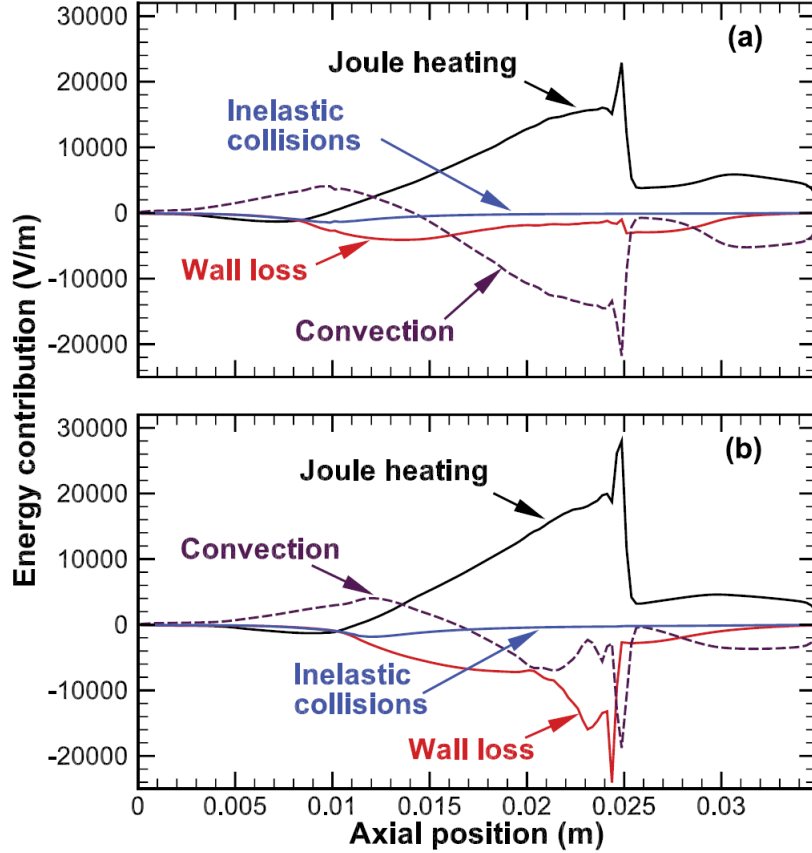


Figure 2.11. Simulations show that Joule heating is balanced mostly by convection in the breathing mode (a) but by wall loss in the stable mode (b) [55]. Reproduced from Figure 11.

Hara et al. [56] further investigated the stability conditions of low-frequency oscillations. He pointed out that the oscillation amplitude of the predator-prey equations is constant and determined only by the initial conditions, and so cannot be used to study stability conditions. A 0D analytical model with state variables (densities, temperatures, velocities) averaged over the discharge channel can be written as:

$$\begin{aligned}
 \frac{\partial n_i}{\partial t} + n_i \frac{V_i}{L_{ch}} + \frac{2n_i U_{i,w}}{R_\Delta} &= \xi(T_e) n_i n_n \\
 \frac{\partial n_n}{\partial t} + (n_n - n_{int}) \frac{V_n}{L_{ch}} &= -\xi(T_e) n_i n_n,
 \end{aligned}
 \tag{2.7}$$

where n_{int} is the neutral density at the anode flowing into the discharge channel, $U_{i,w} = (eT_e/M_i)^{1/2}$ is the ion acoustic speed, and R_Δ is the channel width. This model is similar to the original Fife equations, but note that it is more correct with a constant incoming neutral flow from n_{int} , with ion diffusion to the channel walls included, and with the ambiguous ionization length L replaced

with the channel length L_{ch} . The model considers parameters averaged over the length of the channel.

Hara showed that oscillations of these equations are unconditionally stable (always damped) if the electron temperature is constant because the constant neutral flow damps oscillations. They can be made unstable if the ionization rate constant, or equivalently electron temperature, oscillates. Simulations showed that the oscillation in ion density is still damped when the ionization rate constant has an oscillation amplitude of 0.1%. Surprisingly, a linear mode with an ion density oscillation amplitude about 50% of the average resulted from an oscillation amplitude of only 1% in the ionization rate. Surprisingly, an oscillation amplitude of only 1% in the ionization rate resulted in a linear mode with oscillation amplitude of 50% of the average value in ion density. A nonlinear mode with amplitude much larger than the average ion density resulted from only 5% peak-to-peak amplitude in ionization rate (cf. Figure 10 in Ref. [57]). It was thus argued that an oscillation in electron temperature could determine the stability/excitation of low-frequency oscillations. Note that oscillations in electron temperature are known to be associated with the breathing mode. Lobbia measured an electron temperature oscillation in both the BHT-600 [7] and X2 [58] Hall thrusters well over the 1% level of relative oscillation amplitude.

To further investigate the effect of electron temperature, an electron energy equation and ion momentum equation were added to the system of equations. It is difficult to derive analytical solutions to the complete system of equations, but numerical simulations showed that the growth rate generally increased if either discharge voltage or anode mass flow rate were increased, a result that he noted was in qualitative agreement with the experimental findings of Sekerak that the stable "local" mode occurred in a narrower range of magnetic field at higher discharge voltage and anode flow rate.

Chapter 3

Laser-Induced Fluorescence

When God said “Let there be light” he surely must have meant perfectly coherent light.

– Charles Townes

This chapter contains the necessary information to understand laser-induced fluorescence measurements. Section 3.1 discusses the basic principles of how LIF measurements work in terms of pumping a transition with the laser and collecting fluorescence light. Second, Section 3.1 touches on how a low signal-to-noise ratio requires typical LIF experimental setups to average-out time-resolved information. Section 3.2 describes the details of line broadening mechanisms. Doppler broadening is the effect enabling measurement of the velocity distribution and as such is the most important subsection. Other broadening mechanisms tend to obscure the VDF in the LIF profile, but fortunately they are usually negligible for the conditions of this dissertation and do not necessarily need to be accounted for (see Subsection 3.2.5). Finally, Section 3.3 introduces the other time-resolved LIF techniques that were developed a few years before or concurrently with the work of this dissertation. A basic description of how they work, their technical capabilities, and some key results are presented as an argument motivating the development of a new time-resolved LIF technique.

3.1 Principles of LIF Measurements

Producing laser-induced fluorescence (LIF) is a two-step process. The laser excites an electronic transition in atoms or ions of interest from a well populated state (typically a metastable or ground state is chosen) to some upper state. By spontaneous emission, the excited atoms then decay to a lower state in a characteristic time on the order of nanoseconds and emit a photon with energy equal to the energy difference between the two states. This process of absorption and subsequent emission of light is called fluorescence.

The particular transitions used are commonly called the LIF scheme. The scheme is called resonant if the same transition is used for both the laser and fluorescence, or called non-

resonant if different transitions (and therefore different wavelengths) are used. Non-resonant LIF is generally desirable to avoid possible interference from collected laser light.

In this dissertation, we use a common three-level system for LIF experiments on singly-charged xenon (Xe II) in Hall thrusters. The laser excites the $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition at 834.7 nm (in air). Ions in the upper level quickly decay via the $6s^2P_{3/2} - 6p^2D_{5/2}^0$ transition, fluorescing at 541.9 nm (in air). The laser-excited transition is highlighted in red and the fluorescence transition is highlighted in green in a partial Grotrian diagram of Xe II energy levels and transitions in Figure 3.1. The lifetime of the upper state is between about 7 and 9 ns [59]. The lifetime is of particular interest for time-resolved measurements because the excitation and fluorescence processes set a fundamental limit on the time resolution possible with a particular LIF scheme. The transitions used are attractive because (1) the transition is easily accessible by diode lasers, (2) there is a high branching ratio to the 541.9 nm transition, and (3) the line is narrow enough to allow a reasonable resolution of the VDF (see the discussion of hyperfine and isotopic splitting in Subsection 3.2.2 and the lineshape in Subsection 3.2.5).

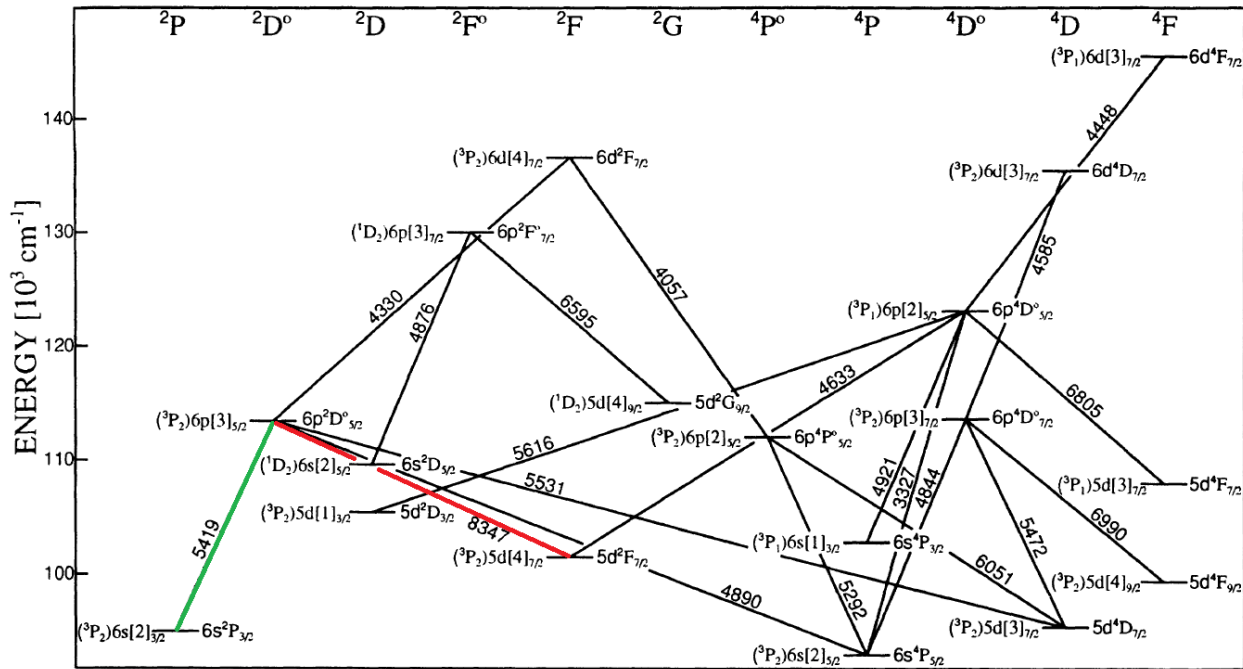


Figure 3.1. Partial Grotrian diagram of Xe II energy levels and relatively strong transitions between them. The laser transition at 834.7 nm is highlighted in red and the fluorescence transition at 541.9 nm is highlighted in green. Note energy is in wavenumbers relative to the Xe II ground state. Modified from FIG. 2 of reference [60].

If not in saturation, then the LIF signal is approximately proportional to the laser intensity and to the density of ions in the lower state of the laser transition (see Subsection 3.2.4 and Equations (3.14) and (3.15)). The background light from spontaneous emission excited by electron impact excitation typically dominates the laser-induced fluorescence signal (see Subsection 5.5.2). Collected light is filtered within some wavelength band by a monochromator or other band-pass filter to remove the background from lines far from the laser transition. The pass band must be wide enough to pass the Doppler shifted fluorescence without distortion from variations in gain as a function of wavelength. Background still dominates even after wavelength filtering within about 1 nm, necessitating some method of signal recovery.

A typical time-averaged experiment modulates the beam with an optical chopper at a frequency on the order of kilohertz, thereby modulating the LIF signal. The modulated signal amplitude can then be recovered from the noise using a lock-in amplifier. The lock-in amplifier averages over a length of time proportional to the time constant setting, which is typically between 100 ms and several seconds for the conditions of an LIF experiment in a Hall thruster. Time-resolved information about oscillations on the order of kilohertz or more is necessarily lost.

3.2 Line Broadening Mechanisms

If the Doppler-free lineshape were a Dirac delta function, the measured LIF profile would correspond exactly to the velocity distribution function (VDF) due to Doppler broadening. In reality, there are many other broadening mechanisms that can significantly influence the LIF profile and may need to be taken into account. The unsaturated LIF profile is a convolution of the Doppler lineshape with all other line broadening mechanisms (e.g. lifetime broadening, hyperfine structure, isotopic shifts) as well as with the laser frequency profile [46], [61]–[64]:

$$\varphi_{LIF}(\nu) \propto a(\nu) \otimes \varphi_D(\nu) \otimes \varphi_L(\nu), \quad (3.1)$$

where $\varphi_{LIF}(\nu)$ is the LIF profile as a function of frequency, $a(\nu)$ is the Doppler-free lineshape, sometimes known as the point spread function or the kernel, $\varphi_D(\nu)$ is the Doppler lineshape, and $\varphi_L(\nu)$ is the laser frequency profile. The Doppler-free lineshape $a(\nu)$ is itself a convolution of the contributions primarily from hyperfine and isotopic splitting and natural broadening [62]. The laser frequency profile has a negligible width of about 1 MHz. The \otimes symbol denotes convolution, a mathematical operation on two functions that generally “mixes” the two functions (see Equation (4.14)).

In general, the Doppler lineshape must be recovered from the LIF profile by deconvolving the other broadening processes out of the LIF profile [62], [65]. In practice, the convolution of the VDF with the kernel representing other broadening mechanisms significantly broadens the LIF profile only if the width of the kernel is relatively close to the width of the LIF profile. This is seldom the case since Doppler broadening tends to be the dominant broadening mechanism of the LIF profile, as will be discussed in the following subsections.

The following subsections cover the main sources of line broadening that are relevant to this dissertation. Doppler shift and broadening is the effect that we exploit to measure the velocity distribution (Subsection 3.2.1). Hyperfine and isotopic splitting (Subsection 3.2.2) and lifetime broadening (Subsection 3.2.3) are the main fundamental sources of the line broadening. If the laser power is set too high, the population will saturate and power broadening will distort the line (Subsection 3.2.4). Finally, Subsection 3.2.5 shows empirical measurements of the lineshape and argues that deconvolution is not necessary for the purposes of this dissertation. Pressure broadening and Stark broadening are neglected because they tend to be important at higher heavy particle density and electron density [66], [67]. Huang reviewed data in the literature and estimated that stark broadening for the Xe II 541.9 nm transition is only 20 MHz [68] for an electron density of 10^{19} m^{-3} , a typical maximum electron density in the H6 Hall thruster [69].

3.2.1 Doppler Shift and Broadening

The Doppler shift leads to a shift in the resonance frequency of the electronic transition excited by the laser if atoms are moving relative to the laser source. For a nonrelativistic ion velocity \vec{v} and a laser of frequency ν_L in the lab frame with a unit wave vector \hat{k} , the change in the laser frequency observed in the ion frame is:

$$\Delta\nu = -\nu_L \frac{\vec{v} \cdot \hat{k}}{c} = -\nu_L \frac{v_k}{c}, \quad (3.2)$$

where v_k is the projection of the velocity along the laser beam and c is the speed of light. Since the laser frequency is approximately the transition frequency, Equation (3.2) implies a handy rule: each gigahertz of Doppler detuning corresponds to a velocity (m/s) of nearly the same numerical value as the transition wavelength (nm).

The frequency of the laser at which excitation and fluorescence occur then gives information about the velocity distribution of the ions via the Doppler shift. If fluorescence occurs at a given laser frequency in the lab frame ν_L , then the Doppler-shifted laser frequency observed in the ion frame equals the zero-velocity transition frequency for the ions that fluoresce:

$$\nu_{trans} = \nu_L + \Delta\nu = \nu_L \left(1 - \frac{v_k}{c} \right), \quad (3.3)$$

where ν_{trans} is the zero-velocity transition frequency. This relationship implies a one-to-one transformation between a given laser frequency and the velocity of a group of ions that correspond to that frequency:

$$\begin{aligned} \nu_L &= \frac{\nu_{trans}}{1 - v_k/c} \\ v_k &= c \left(1 - \frac{\nu_{trans}}{\nu_L} \right) = -c \frac{\Delta\nu}{\nu_L}. \end{aligned} \quad (3.4)$$

Hence, an LIF experiment scans a range of laser frequencies near a particular transition and measures the LIF signal profile as a function of laser frequency. The LIF signal at each frequency is essentially sampling the population of ions at the corresponding velocity. The measured Doppler lineshape in laser frequency $\varphi_D(\nu_L)$ maps directly to the velocity distribution. Assuming the Doppler lineshape $\varphi_D(\nu_L)$ is integral normalized, the velocity distribution is:

$$f(v_k) = \frac{\nu_{trans}}{c \left(1 - v_k/c \right)^2} \varphi_D \left(\frac{\nu_{trans}}{1 - v_k/c} \right). \quad (3.5)$$

The multiplicative factor on the right hand side comes from imposing the normalization that the integral over both distributions equals unity and the change of variable in the integral. Though Equation (3.5) gives the relationship to the formal definition of the 1D velocity distribution, all velocity distributions shown in this dissertation are actually normalized with a peak value of one for ease of interpreting the experimental results.

The relationship between the Doppler shifted/broadened LIF lineshape and the velocity distribution is strongly related to the discussion of Doppler broadening in many spectroscopy texts (e.g. [61], [70]). In the usual development, a Maxwell-Boltzmann velocity distribution is assumed and the resulting broadening of the emission lineshape from the Doppler shift is

calculated. Here we measure the absorption lineshape with LIF and infer the velocity distribution that lead to it, assuming that Doppler broadening is the only broadening mechanism responsible for the measured excitation lineshape.

Note that the Doppler shift is only observed for the projection of velocity in the direction of the laser wave vector, so the LIF profile corresponds to the projection of the ion velocity distribution along the wavevector. In other words, if we define a coordinate system with the laser in the x direction, then the Doppler broadening corresponds only to the one-dimensional velocity distribution:

$$f(v_x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(v_x, v_y, v_z) dy dz, \quad (3.6)$$

where $f(v_x, v_y, v_z)$ is the general velocity distribution from kinetic theory. Orthogonal laser beams can give the 1D velocity distribution in all three dimensions, but the information about how velocity in one dimension is related to velocity in another dimension is lost. Consider, for example, the intersection of two ion beams that together produce two distinct peaks in both x and y velocity projections. With the two 1D VDFs, the information about which peak in the x projection corresponds to the same ions as which peak in the y projection is lost.

This subtlety is important to keep in mind in interpreting LIF data, but is not usually important for electric propulsion thrusters, where there is usually only one ion population at a given point in space and time. Even for special cases where more than one ion population exists, such as the intersection of ions from all sides of the annulus at a Hall thruster centerline, the origin of the main features of the LIF profile can be generally understood without a direct measurement of the general velocity distribution (see the discussion of Fig. 2 in [71]). In that example, however, a measurement of the full velocity distribution might help explain the origin of the ions with low radial speed at the thruster centerline. It is possible to infer the full velocity distribution from LIF data by using tomographic reconstruction, but it requires LIF profiles in more than three directions, and a sophisticated optical setup inside the chamber [72].

3.2.2 Hyperfine and Isotopic Splitting

The most significant contribution to the Doppler-free lineshape comes from hyperfine and isotopic splitting. Xenon gas contains 9 stable isotopes in the abundances listed in Table 3.1. The total nuclear spin is zero (quantum number $I = 0$) for the seven isotopes with an even number of nucleons. The odd-numbered isotopes, however, have a non-zero total nuclear spin

quantum number of $I = 1/2$ for Xe-129 and $I = 3/2$ for Xe-131. The coupling of the total electron spin and total orbital angular momentum gives rise to the fine structure, a splitting of the energy states for different values of the total electron angular momentum quantum number J . Similarly, the electromagnetic coupling of the total nuclear spin and total electron angular momentum leads to hyperfine splitting different states of the total angular momentum quantum number F [73].

Table 3.1. Xenon isotopic abundance in percent from reference [28].

Isotope	124	126	128	129	130	131	132	134	136
Abundance	0.0096	0.0090	1.919	26.44	4.075	21.18	26.89	10.44	8.87

The total angular momentum quantum number F can take values in the range:

$$F = I + J, I + J - 1, \dots, |I - J| + 1, |I - J|. \quad (3.7)$$

The selection rule for allowable transitions between upper and lower states with total angular momentum quantum numbers F_U and F_L is that F must change by at most 1 up or down. The one exception is that a transition between states with $F_U = F_L = 0$ is forbidden:

$$\Delta F = F_U - F_L = [0, \pm 1], F_U = 0 \nrightarrow F_L = 0. \quad (3.8)$$

The shift in energy due to nuclear spin coupling is given by the following [73]:

$$E_{hfs} = A \frac{C}{2} + BD$$

$$C = F(F + 1) - I(I - 1) - J(J + 1) \quad (3.9)$$

$$D = \frac{(3C/4)(C + 1) - I(I + 1)J(J + 1)}{2I(2I - 1)J(2J - 1)},$$

where C and D are functions of the quantum numbers of the states and A and B are known as the magnetic dipole and electric quadrupole constants, which are different for each transition and isotope. For the $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition, this energy splitting gives rise to 3 distinct components for Xe-129 and 9 for Xe-131.

The seven even-numbered isotopes, having no total nuclear spin ($I = 0$), do not have the splitting described above from nuclear spin coupling. Differences in the mass and volume of the different isotopes, however, do lead to line splitting among the even-numbered isotopes and a further shift that is the same for the line components of each odd-numbered isotope. Then there are a total of 19 different line components for the $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition. The line

splitting can be modeled as a series of Dirac delta functions as shown on the left frame of Figure 3.3. The “normal mass effect” is due to movement of the nucleus since the nucleus is not infinitely heavy relative to electrons [73]. The “specific mass effect” is due to correlations between outer electrons. Mass effects tend to dominate for light atoms, while volume effects dominate for heavy atoms. With moderate weight, xenon is affected by both. The volume effect occurs because heavier atoms have a larger nuclear radius, into which the electron charge distribution can penetrate, changing the interaction from just the Coulomb field. Each isotope and transition therefore has a different isotopic shift. These effects can be on the same order of magnitude as hyperfine splitting. Calculating isotopic shift is difficult for complicated atoms like xenon, and the shift is most often measured experimentally.

The relative intensity of line components for the even isotopes is simply proportional to the relative abundance of the isotopes. The relative intensity of the nuclear spin-splitting line components is also proportional to the abundance of the isotope to which they belong, as well as to the following factors for a $J \rightarrow J - 1$ transition [74]:

$$\begin{aligned}
 I(F \rightarrow F - 1) &\propto \frac{P(F)P(F - 1)}{F} \\
 I(F \rightarrow F) &\propto \frac{(2F + 1)}{F(F + 1)} P(F)Q(F) \\
 I(F - 1 \rightarrow F) &\propto \frac{Q(F)Q(F - 1)}{F} \\
 P(F) &= (F + J)(F + J + 1) - I(I + 1) \\
 Q(F) &= I(I + 1) - (F - J)(F - J + 1).
 \end{aligned} \tag{3.10}$$

If the hyperfine splitting constants are known, then the contribution to the lineshape from hyperfine splitting can be modeled and deconvolved out of the LIF profile. Unfortunately, the constants are not full known for the $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition. If unknown, then the hyperfine contribution to the lineshape can be estimated from measurements as shown in Subsection 3.2.5.

3.2.3 Lifetime Broadening

Lifetime broadening, or natural broadening, is one mechanism that ensures that all atomic lines have finite width. It can be intuitively understood by considering the Heisenbug uncertainty principle in quantum mechanics. Just as position and momentum are complementary observables

that cannot together be known to an arbitrary accuracy, so are energy and time. The product of their uncertainties has a lower bound:

$$\Delta E \Delta t \geq \frac{\hbar}{2}, \quad (3.11)$$

where ΔE is the uncertainty in energy and Δt is the uncertainty in the time interval during which energy is measured (the standard deviations of the distributions). For stationary states with small relative energy uncertainty, the equality holds. The exponential distribution, which describes the probability distribution of the time between events, has uncertainty equal to the average:

$$\Delta t = \tau_p, \quad (3.12)$$

where τ_p is the radiative lifetime of the state p , which is the inverse of the sum of all rate constants (Einstein coefficients) for transitions out of state p . Because the energy of the transition is the difference between the initial and final energy levels, the uncertainty in the photon energy depends on the lifetime of both the upper and lower states. The lineshape from this process as a function of angular frequency for a transition from state p to state q is Lorentzian with a half width of $\Delta\omega$ where [70]:

$$I(\omega) = \frac{\Delta\omega}{(\omega - \omega_0)^2 + (\Delta\omega/2)^2} \quad (3.13)$$

$$\Delta\omega = \frac{1}{\tau_p} + \frac{1}{\tau_q}.$$

The lifetime of the upper state is about 7 to 9 ns [59], and the lower state is metastable and can be considered infinite relative to the lifetime of the upper state. Taking the lower bound for a worst case estimate of the natural linewidth in ordinary frequency results in $\Delta\nu = 23$ MHz, or about 19 m/s in velocity. Lifetime broadening is therefore negligible for the typical LIF measurement in a Hall thruster.

3.2.4 Fluorescence Signal Saturation and Power Broadening

At low laser power, the fluorescence signal is proportional to the input laser power. This trend does not continue indefinitely; the fluorescence signal eventually reaches an asymptotic limit and becomes independent of the laser power. More importantly for LIF measurements of the VDF, this saturation also causes an apparent broadening of the spectral lineshape, called saturation broadening or power broadening, which is another influence to the LIF profile that may need to be removed to measure the VDF. As we will see, Power Broadening is a subtle

effect that is difficult to fully account for in general, and therefore it is usually preferable to avoid saturation rather than treat it.

The concept of saturation can be readily understood by considering a two-level system. Starting from a simplified rate equation for the density of the upper state population, we have [61]:

$$\frac{dn_2}{dt} = B_{12}I_0n_1(t) - (B_{21}I_0 + A_{21} + Q_{21})n_2, \quad (3.14)$$

where B_{12} , B_{21} , A_{21} are the Einstein coefficients for photon absorption, stimulated emission, and spontaneous emission for the transition between the states 1 and 2, Q_{21} is the quenching rate constant, I_0 is the laser intensity, and $n_1(t)$ is the lower state density. With the simplifying assumption of a constant number of particles $n_0 = n_1(t) + n_2(t)$, the steady-state solution is:

$$n_2 = \frac{B_{12}I_0n_0}{A_{21} + Q_{21} + (B_{21} + B_{12})I_0}. \quad (3.15)$$

Notice that at low intensity, specifically when the quenching and spontaneous emission constants dominate the denominator, the upper state population is proportional to laser intensity. At high intensity, when the term with photon absorption and stimulated emission dominate the denominator, the upper state population (and therefore fluorescence intensity) is invariant to laser intensity.

The saturation intensity is defined as the laser intensity corresponding to an upper-state population or fluorescence intensity of half of the limiting value. The above equation can be rewritten in the form of:

$$n_2 = \frac{n_0}{1 + \frac{g_1}{g_2}} \frac{I_0 / I_{sat}}{1 + I_0 / I_{sat}} \quad (3.16)$$

$$I_{sat} = \frac{A_{21} + Q_{21}}{B_{12} \left(1 + \frac{g_1}{g_2}\right)},$$

where I_{sat} is the saturation intensity and g_1 and g_2 are the degeneracies of the two levels. The dimensionless ratio of laser power to saturation intensity is called the saturation parameter. The saturation intensity can depend not only on constants but also on the conditions of the experiment via the quenching rate. The definition of saturation intensity in Equation (3.16) is intuitive because the spontaneous emission and quenching rate constants make it harder for the system to reach the ultimate upper-state density. Conversely, the photon absorption rate constant

makes it easier, reducing the laser intensity necessary for saturation. It can be shown by an argument equating the rate equation formalism to collisional cross section formalism that the saturation intensity essentially corresponds to one photon per ion per lifetime of the upper state [75]. In other words, laser power near and above the saturation intensity is less efficient at producing LIF because many photons are incident on already-excited ions.

In measuring the LIF profile, the narrow-bandwidth laser (~ 1 MHz) is applied to the line not only at the peak but sequentially swept across the lineshape. The laser will be less effectively absorbed and not saturate the population as easily at the edge of the lineshape than at the center. Hence, for a narrow-bandwidth laser applied at some arbitrary frequency, the saturation intensity actually varies as a function of laser frequency inversely proportionally to the lineshape [62], [75]. As laser power rises, the peak of the lineshape will saturate first and stop responding as strongly to the laser, and therefore the edges of the LIF profile will have disproportionately high signal, leading to an apparent broadening of the line.

The specifics of the broadening depend on the unbroadened lineshape. The traditional treatment of power broadening for a Lorentzian lineshape transforms it into another Lorentzian function with the width broadened by [75]:

$$\Delta\omega_b = \sqrt{1 + I_0 / I_{sat}} \Delta\omega_a, \quad (3.17)$$

where $\Delta\omega_b$ is the broadened width and $\Delta\omega_a$ is the unsaturated width. Smith developed a method to model the power broadening for an arbitrary lineshape and to measure the unbroadened lineshape [62]. Figure 3.2 shows the simulated power broadening of the Xe II $5d^4D_{7/2} - 6p^4P_{5/2}^0$ transition.

Huang pointed out that removing the effect of saturation broadening from the LIF profile can be very complicated because the saturation parameter will vary with the laser intensity profile in space [68]. If it is a Gaussian beam, then the center of the beam/interrogation volume will have a higher saturation parameter than at the edges, and removing the effects of saturation broadening would require detailed measurements of the beam profile and interrogation volume. He concluded that it would be best to avoid saturation rather than treat it.

Fortunately, saturation is not typically a problem for the laser power, beam waist, transition, and vacuum facility used in the experiments of this dissertation. Beam power is often about 10 mW at the interrogation volume with a beam diameter about 1 mm. Saturation was

checked during some of the cathode experiments of this dissertation by measuring the peak signal level as a function of laser power using neutral density filters. The plot of fluorescence signal versus laser power was in the linear regime. The saturation test data are not presented here because it is not important for the cathode campaign since it was solely for validating the system and the actual lineshape was not of interest. Unfortunately, saturation tests only provide an ad hoc check that would have to be redone with each experiment unless carefully calibrating for changes to the laser power in the chamber, the beam size/shape, and the background pressure. Time did not allow for saturation tests with the H6 campaign, but Huang et al. [45] performed a saturation study with the same thruster, vacuum facility and similar optics as this dissertation and did not detect saturation.

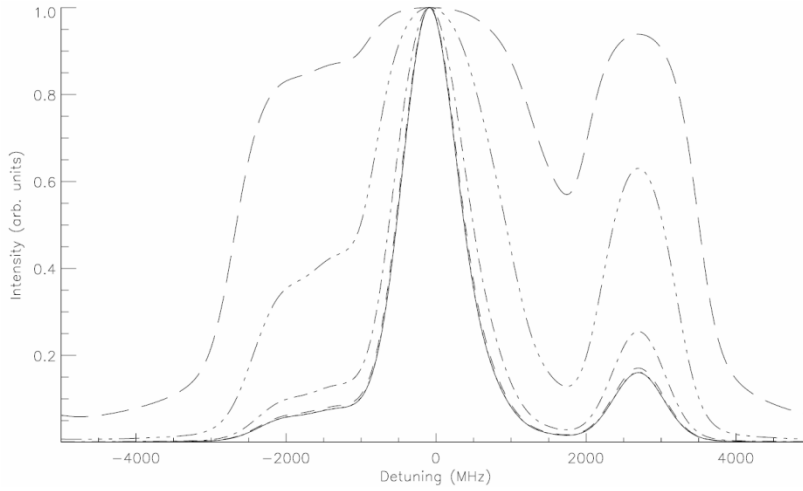


Figure 3.2. A simulation of the power broadening of the $\text{Xe II } 5d^4D_{7/2} - 6p^4P_{5/2}^0$ transition for a variety of laser power levels. The first profile where the side peak is noticeably higher corresponds to $\max(I_0/I_{sat}(\nu)) = 0.79$, and the first profile with a significantly broadened main peak is at $\max(I_0/I_{sat}(\nu)) = 7.92$. Reproduced from [62].

3.2.5 Lineshape of the Transition

With known values for isotopic shift and hyperfine splitting constants, the kernel can be modeled and deconvolved out of the LIF profile by a method such as developed by Smith [62]. Alternatively, the kernel can be measured empirically with a source whose VDF shows only minor Doppler broadening [46] or, more accurately, by a Doppler-free method [65]. The hyperfine constants for the $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition were not known until recently, and measurements are still not entirely accurate or complete, particularly for the isotopic splitting [67]. The difficulty in measuring the constants for this line is that many line components

completely overlap under only natural broadening and slight Doppler broadening. The FWHM of a slightly Doppler broadened line used to calculate hyperfine constants was about 750 MHz (left panel in Figure 3.3) [67], while the FWHM of the Doppler-free lineshape from another study was 440 MHz (right panel in Figure 3.3) [65]. These widths correspond to 626 m/s and 367 m/s in speed.

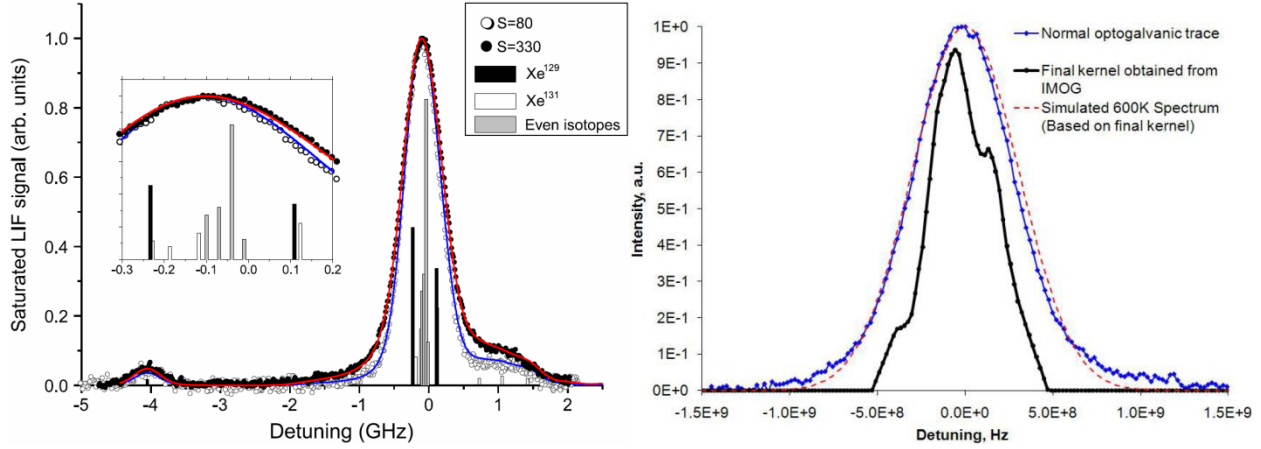


Figure 3.3. Left: A slightly Doppler broadened lineshape of the $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition has a FWHM of 750 MHz (reproduced from [67]). Right: the Doppler-free lineshape measured in another study has a FWHM of 440 MHz (reproduced from [65]). The widths are only 626 m/s and 367 m/s in speed.

Though the narrow lineshape confounds measurement of the hyperfine splitting and isotopic shift constants; on the other hand, it is fortunate that the narrow lineshape rarely significantly affects the VDF measurement. Conceptually, if the kernel is much narrower than the VDF then it can be regarded approximately as a Dirac delta function, and the convolution of the VDF with a delta function is the VDF (see the definition of convolution in Equation (4.14)). The width of the LIF profile is normally well over 1 GHz (a few km/s in speed) at most locations. Convolution of such a wide VDF with a kernel of a width about 440 MHz may smooth out fine features but does not strongly affect the LIF profile or its width.

This is true especially since the widths of convolved functions do not generally add directly (though they do add for the special case of Lorentzian functions [70]). For example, the convolution of two Gaussian functions with widths $FWHM_{G1}$ and $FWHM_{G2}$ is another Gaussian with the widths added in quadrature [70]:

$$FWHM_{G1 \otimes G2}^2 = FWHM_{G1}^2 + FWHM_{G2}^2. \quad (3.18)$$

For the example of a very narrow Doppler profile with a width of only 1 GHz and a kernel with a width of 440 MHz, the resulting LIF profile is broadened by only 93 MHz if both are Gaussian. In other words, the width of the resulting LIF profile would have an error of less than 10% for a kernel that is 44% as wide as the Doppler profile (again, assuming that both are Gaussian). This error is likely insignificant for most experiments, especially for a time-resolved measurement that may have worse SNR than a time-averaged measurement and resolve only about 20 points in velocity space. Even for a very narrow profile with a width of 1 GHz, 20 points implies a step size of about 50 MHz, almost the magnitude of the broadening. For a more realistic LIF profile with a larger width, not only would the broadening be reduced, but the step size between points in velocity space would increase, rendering the error undetectable.

Similarly, Huang et al. [65] concluded that the deconvolution is only significant with the H6 for the very narrow LIF profiles found upstream of most of the acceleration and ionization. Even there, the difference in the FWHM was less than 10%, though he noted that the error in the estimate of temperature could be more than 10%. Hargus and Nakles [46] also found that deconvolving an empirically measured kernel is insignificant for the BHT-200-X3. Since Doppler broadening alone is the dominant contribution to the LIF profile, some authors using this transition report the LIF profile directly as the VDF without deconvolving the kernel corresponding to other sources of broadening and splitting [46], [64], [71], [76].

No deconvolution has been performed in this dissertation; all VDFs are the raw LIF profiles. This was a practical decision to avoid overcomplicating the already complicated signal processing of the time-resolved technique. As discussed above, time-resolved measurements with only about 20 points in velocity space already have a poor resolution in velocity, and deconvolution would not help much. In future work with more points and a better SNR, however, deconvolution of time-resolved LIF profiles may be necessary to accurately resolve features of the VDF since the instantaneous LIF profile can be much narrower than the time-averaged LIF profile.

3.3 Time-Resolved LIF Techniques

The main difficulty in measuring LIF signal in Hall thrusters is the strong background emission that obscures the LIF signal (see Subsection 5.5.2). The background emission makes a direct measurement of LIF signal as a function of time impossible and requires any signal recovery technique to average over a large amount of data collected over many oscillation

periods. An indirect measurement of the time-resolved LIF signal can be possible if the signal adheres to certain assumptions that can be exploited by an averaging scheme. This subsection provides an overview of the time-resolved LIF techniques that have been used with Hall thrusters or related plasma sources to date, as well as a sampling of some key results when relevant to the results of this dissertation.

Some of the advantages and disadvantages of each technique are highlighted in this discussion. In particular, note that these techniques generally have limitations to some form of periodically repeating oscillation due to the need to average over many cycles of the oscillation in some way. This limitation motivates the need for a new TRLIF technique that is capable of recovering the signal from nonperiodic oscillations such as undriven/unperturbed Hall thruster breathing and spoke oscillations despite averaging over many oscillation periods. Though some advantages and disadvantages are mentioned here, a more comprehensive comparison of each TRLIF technique, including the new technique developed in this dissertation, is left for Section 8.1 of the concluding chapter.

3.3.1 Combination of Phase-Sensitive Detection and Triggered Averaging

Perhaps the most straightforward way to modify a time-averaged LIF experiment for time-resolved signal recovery would be to reduce the time constant τ of the lock-in amplifier to smaller than the normal value for a time-averaged experiment, about 100 ms, which sets the time scale over which data are averaged. For reasons we will see in Subsection 4.3.2, the time constant τ cannot be made smaller than the laser modulation period, hence a high-speed modulator such as an acousto-optic modulator (AOM) must be used instead of a mechanical chopper for sub-millisecond time resolution. The signal-to-noise ratio, proportional to $\sqrt{\tau}$, will quickly become unacceptable as the time constant is reduced, which is why a large time-constant is used in the first place. One way to keep a small time constant but improve SNR to an acceptable level would be to use a second form of averaging in addition of the phase-sensitive detection (PSD) of the lock-in amplifier. For example, an oscilloscope could be set to average over a large number of waveforms output from the lock-in amplifier triggered based off of some reference signal such as the phase of the discharge current oscillation.

The straightforward technique described is essentially the technique developed by Scime *et al.* for use with pulsed plasma sources [77]. They implemented it using only the equipment of a normal, time-averaged LIF experiment: a commercial lock-in amplifier, oscilloscope, and a

mechanical chopper. The use of the chopper limited time resolution to about 1 ms, and the system was later upgraded with an AOM to achieve a better time resolution [78] on the order of tens of microseconds.

This general technique of combining phase-sensitive detection with a triggered average can also be achieved by digitizing raw data from the photomultiplier and applying the PSD and triggered average in software in post-processing. That is significant because it is the method of the "triggered averaging" technique that will often be used in this dissertation to compare with the transfer function averaging method.

One advantage of this system is that the phase-sensitive detection that demodulates the signal acts as a narrow band-pass filter that rejects much of the noise spectral density before the primary averaging process (see Subsections 4.3.1 and 4.3.2). This allows for a trade-off between time-resolution and SNR tuned to an appropriate value for the experiment. It generally has a shorter acquisition time than techniques that use a triggered average more directly as the only form of signal recovery, such as the synchronized photon counting of the following subsection. Biloiu, reports averaging over only "a few hundred" plasma pulses, but the results for a pulsed RF plasma source with LIF operating on an argon line are completely incomparable with the results for Hall thrusters operating on a Xenon line. The triggered averaging reported in this dissertation for a single-wavelength, 60-s-long dataset averages over about 50 thousand triggered waveforms.

The main disadvantage of the technique is that it implicitly assumes a repeatable process occurs after trigger so that averaging the triggered waveforms together will remove noise and recover the signal waveform (see Subsection 4.3.4). While this may be a very reliable assumption for pulsed plasma sources, it is not an accurate assumption for most Hall thruster oscillations. If triggered averaging is applied to nonperiodic oscillations with a varying period, the resulting waveform unphysically decays as the many time series begin in phase at the trigger but drift out of phase from each other in time (see Section 6.4).

3.3.2 Synchronized Photon Counting

The first technique used to measure time-resolved LIF in Hall thrusters was pioneered by Mazouffre et al. It involves low speed laser modulation (~ 20 Hz) and a photon counting technique that discriminates signal counts from noise by adding (subtracting) counts that were collected while the laser is on (off) . The system triggers the averaging at a given phase of the

oscillation and counts are added to and subtracted from the time bins over many trigger cycles. Hence this technique essentially performs a triggered average over an ensemble of waveforms to recover an ensemble average waveform, but it does so over photon counts instead of a digitized voltage signal.

Since low-frequency Hall thruster oscillations under normal operating conditions are nonperiodic, all tests using this technique on Hall thrusters have perturbed the operating conditions to force oscillations more amenable to the triggered average. In the initial series of papers [79]–[81], the discharge current was periodically cutoff to create a high-amplitude, repeatable oscillation (the reignition of the thruster). Periodically cutting the discharge current was found to change the average ion VDF [80]. More recently, a low-amplitude, sinusoidal perturbation to the cathode keeper potential has been used [82]–[84], which does not change the average ion VDF or average plasma parameters [82]. The power spectral density of the discharge current is mostly unchanged except for the introduction of a very sharp peak at the driving frequency. The discharge current oscillation amplitude is increased. It is unknown how the perturbation changes the time-resolved LIF signal other than simply forcing a more periodic oscillation and increasing amplitude.

The tests using a periodic cutoff of the discharge current found an oscillating mean velocity near the exit plane and a front of electric field (calculated from the measured ion VDF) propagating from outside the exit plane to slightly upstream [80], [81]. Because the velocity spread remained approximately constant, they suggested the ionization zone and acceleration zone (E-field) may oscillate together axially in phase, a finding consistent with some simulations [52]. The velocity of the front is found to be approximately the neutral speed, taken as evidence of the importance of the neutral gas temperature.

The discharge stabilization with the keeper potential enabled interrogation of high-frequency (100 kHz - 1 MHz) oscillations in the ion VDF and electric field [83] and comparison of the observed oscillations with simulations of ion transit time oscillations [84]. Ion transit oscillations were predicted and found only outside the exit plane in the negative magnetic field gradient region. An example ion VDF measured near the exit plane of a 200-W Hall thruster is shown in Figure 3.4. High-frequency oscillations were detected in an average over many cycles while triggering the averaging based on the low-frequency driving signal of the cathode. This

implies that the high-frequency oscillations are controlled by and coherent with low-frequency oscillations.

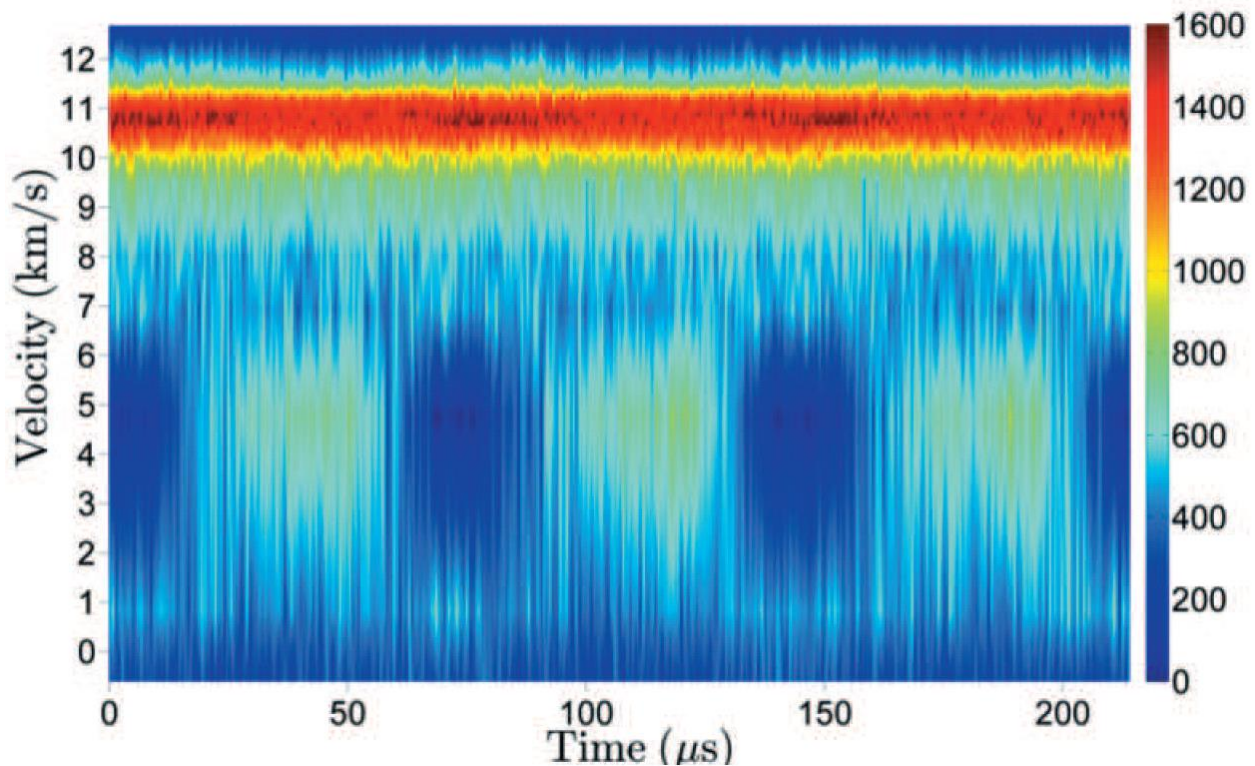


Figure 3.4. Ion VDF measured 4 mm downstream of the exit plane of a 200-W permanent magnet Hall thruster [84]. Reproduced from Fig 5.

The main advantage of the synchronized photon counting technique is the superior time resolution, which is limited by the bin size of the multichannel scaler (100 ns for the references in this subsection). Observation of the high-frequency oscillations mentioned above may not necessarily be beyond the capabilities of other techniques, but the above publications are the first to report on such high-frequency oscillations in LIF measurements in Hall thrusters.

One disadvantage is the longer acquisition time necessary for a reasonable signal-to-noise ratio, on the order of many minutes per laser wavelength. The earlier experiments with the discharge current cutoff reported averaging over 1 million cycles, and the later experiments with the keeper oscillation reported "hundreds of thousands." The same xenon line is used, but the plasma sources and optics are completely different, so this is not directly comparable with the number from triggered averaging as a fundamental measure of the averaging efficiency of the techniques. Nonetheless, the order-of-magnitude difference in the number of triggers averaged does stand out and is likely partially due to the averaging techniques. The improvement factor for

the signal-to-noise ratio is analyzed more rigorously in Appendix B, which confirms that the SNR rises as the square root of the number of triggers averaged and the square root of the time bin size.

The necessity to perturb the natural oscillations to create periodic oscillations amenable to the triggered average is another main disadvantage. Since this technique is essentially a different implementation of the triggered average, the key implicit assumption is that a repeatable event occurs at each trigger. Synchronized photon counting certainly is a powerful technique that has so far shed more light on Hall thruster ion dynamics than any other TRLIF technique. Even so, synchronized photon counting as it has been implemented will not be able to recover the TRLIF signal for unperturbed, nonperiodic Hall thruster oscillations.

3.3.3 Heterodyne TRLIF Signal Recovery

A heterodyne technique was developed by Diallo *et al.* during the time of this dissertation work [11], [85]. It does not involve a triggered average, but nonetheless the key assumption of the heterodyne technique is that the oscillation is periodic at some frequency. An arbitrary periodic signal can be decomposed into a Fourier series that includes components only from the oscillation frequency and its harmonics. The heterodyne technique involves injecting the laser at some modulation frequency and then using a normal lock-in amplifier to recover the heterodyne signal at a frequency of $n\omega_D \pm \omega_L$, where ω_D is the driving frequency (or natural frequency of at least quasi-periodic oscillations), ω_L is the laser modulation frequency, and n is the order of the frequency component collected.

In the reported experiments to date, the anode potential was driven at the peak frequency of the natural breathing oscillations, forcing more periodic oscillations. The system was able to reconstruct the time-resolved ion VDF based on the zero and first order components. The results generally bear a resemblance to ion VDFs measured with the synchronized photon counting technique and the measurements of this dissertation in Chapter 7. Figure 3.5 shows an example of a reconstructed VDF at the exit plane of a cylindrical Hall thruster (see Subsection 2.3.4). Similar to the low-field condition of the H6, there is a broad oscillation with a weak low-energy tail, and the signal briefly collapses to the noise floor.

Reconstructed time-dependent IVDF

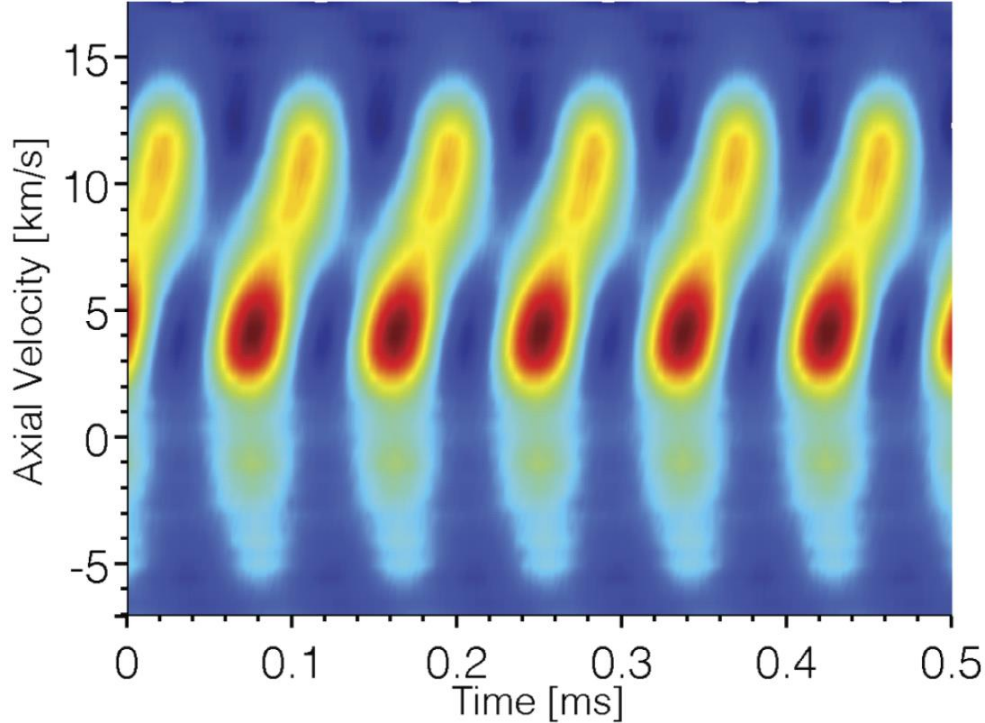


Figure 3.5. The time-resolved ion VDF at the exit plane of a cylindrical Hall thruster, reconstructed to second order in the harmonic decomposition [11]. Reproduced from FIG. 4.

One advantage of this technique is that it can be performed with essentially the same equipment as a standard time-averaged LIF experiment in a slightly different configuration. The laser is modulated at a low frequency with a chopper, and a lock-in amplifier recovers the heterodyne signal of the Fourier frequency components instead of the time-averaged RMS signal amplitude at the frequency ω_L .

Another advantage is that the Fourier series components for each laser wavelength can be measured in about the same acquisition time as the time-averaged signal of a normal experiment. The entire VDF scan was reportedly done in under 10 minutes. The acquisition per laser wavelength is therefore fast enough to allow very good resolution in velocity space with hundreds of points. This compares very favorably with most other TRLIF techniques, where it is typically only practical to capture 10 to 20 points.

A significant disadvantage, like the other two methods described so far, is that it requires a periodic oscillation for the assumption of decomposing the signal into a Fourier series. Therefore, the heterodyne technique will not be appropriate for many natural Hall thruster

oscillations, though it may be possible to acquire data at some quasi-periodic operating conditions.

In addition, the reconstruction is inherently limited to some maximum measurable harmonic order. This limitation raises questions about the accuracy of the low-order reconstructed waveform. The accuracy of the low-order reconstruction is unknown even if the assumption of a periodic oscillation is valid for the driven oscillation and the low-order Fourier series components are measured to perfect accuracy. The technique was originally demonstrated to first order due to noise issues in recovering higher order signal [85]. It was later demonstrated to second order, with the second order component apparently giving only a small contribution for the sinusoidal driving waveform [11].

This limitation means that not only is the accuracy of the waveform of each velocity group unknown, but the actual time resolution of the system is unclear. If the reconstruction is a partial Fourier series that includes only the fundamental frequency and a small number of harmonics, what is the smallest feature in time that can be accurately resolved? If a small feature is resolved, will it introduce ringing in the rest of the time series? The smoothness in velocity space is an impressive result of the large number of different laser wavelengths used for the reconstruction. The smoothness in time, however, should be taken with some skepticism because it is an artifact of reconstructing the waveform as a sum of sine waves, a continuous function that can be sampled with an arbitrary sampling frequency.

3.3.4 *Sample-And-Hold TRLIF Signal Recovery*

MacDonald, *et al.* developed a TRLIF system that is "tolerant of natural drifting in the current oscillation frequency." The system, developed around the same time as much of this dissertation work [86], uses a voltage comparator to trigger at a consistent discharge current level, approximately corresponding to a given phase of the discharge current oscillation. A delay generator and a gated sample-and-hold circuit then hold the signal level at a given delay time after each trigger. The output from the sample-and-hold circuit, updated at each trigger, is sent to a lock-in amplifier to recover the signal from the noise. Note that the laser is modulated at a frequency lower than the discharge oscillations so that the sample-and-hold output is updated many times per modulation period. The modulation frequency was 11 Hz for the 60-Hz oscillations of a xenon lamp in the original paper [86] and 2.5 kHz for the 48-kHz oscillations of the BHT-600 Hall thruster in a more recent paper [87]. Since only the signal during a short gate

at a fixed phase point is input to the lock-in amplifier, the lock-in amplifier's output is the TRLIF signal at the time corresponding to the delay time after the trigger. A typical lock-in amplifier used for time-averaged LIF can recover TRLIF signal at a number of desired delay times to recover a complete waveform. The total acquisition time was reduced by using six separate sample-and-hold circuits and lock-in amplifiers to collect the TRLIF at many delay times simultaneously [87]. Together, they enable recovery of the time-averaged signal and 23 phase points in four laser scans.

Strictly speaking, only the trigger itself is locked to a discharge current level. Each of the phase points after the initial point at the trigger are taken after some fixed delay time from the trigger. Therefore, they are not quite locked to a discharge current level or phase of the oscillation since the period may jitter or drift. In this sense, the technique is strongly related to the triggered average and synchronized photon counting techniques, both of which trigger at some phase of the discharge current oscillation and then average either voltage or photon counts occurring in fixed time bins after the trigger. The gated sample-and-hold circuit, however, is set to average over a time commensurate with the expected jitter in the oscillation period, 1 μs in the case of BHT-600 nominal operating condition, averaging out the jitter/drift and also limiting the time resolution of the system. The averaging effect of the gated sample-and-hold circuit is similar to the averaging effect of phase-sensitive detection in the technique discussed in Subsection 3.3.1 in that they both raise SNR by averaging TRLIF signal over a time scale about 1 microsecond.

An advantage of this technique is that it is more tolerant to acquire quasiperiodic waveforms than a triggered averaging using smaller time bins such as the synchronized photon counting technique. Assuming a constant waveform shape and amplitude but possibly drifting or jittering oscillation period, the comparator triggers at the same phase. Then, as discussed above, the signal is averaged over a gate that is about the size of the jitter so that the signal approximately corresponds to the TRLIF signal at a certain phase of the oscillation.

Another advantage is that the velocity resolution can be nearly the same as a time-averaged experiment. A complete time-series for a single-wavelength is effectively captured in four scans with a time constant of 3 s on the lock-in amplifiers, for a total of 12 s per laser wavelength. This is faster than all other techniques except for the heterodyne technique.

One disadvantage is that this technique resolves a small number of points in time compared with the other techniques and only a single period of oscillation. The time resolution of the scheme is limited by the width of the gate over which data are averaged. A larger number of points could help resolve more than one period, but acquiring data at longer delay times from the trigger would likely result in the typical dephasing of a triggered average over nonperiodic waveforms.

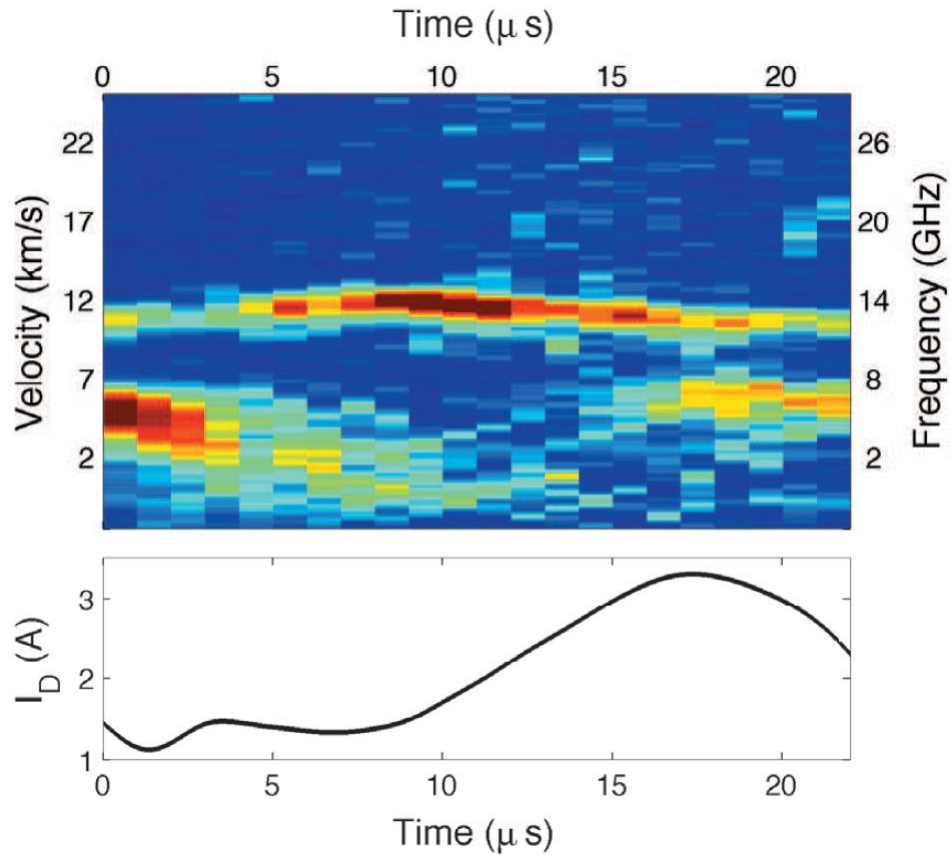


Figure 3.6. Axial time-resolved ion VDF of the BHT-600 at 15 mm downstream of the exit plane on the thruster centerline [87]. Reproduced from Figure 8.

Chapter 4

The Theory of the TFLIF Technique

Experience without theory is blind, but theory without experience is mere intellectual play.

– Immanuel Kant

Armed with the overview of the problem in Chapter 1, the background regarding EP in Chapter 2, as well as further background on LIF in Chapter 3, we are now in a position to begin describing a solution to the problem. This chapter introduces the theory of the transfer function laser-induced fluorescence (TFLIF) technique, the solution that was developed over the course of this dissertation research.

First, Section 4.1 defines the terms used in the dissertation that could be misunderstood. To provide context for how the measurement is made, Section 4.2 outlines the equipment necessary for a TFLIF measurement, though many of the technical details regarding the hardware are left for Section 5.1. Finally, Section 4.3 details the signal processing theory underlying TFLIF measurements: band-pass filtering (Subsection 4.3.1), phase-sensitive detection (Subsection 4.3.2), and Fourier analysis (Subsection 4.3.3). Subsection 4.3.4 describes triggered averaging as an alternate averaging method to use in lieu of the Fourier analysis at the heart of the TFLIF technique. The alternate method is useful as a benchmark to demonstrate that the Fourier analysis converges to the expected result in certain special cases.

4.1 Definitions and Conventions

To avoid possible confusion, this section explains the meaning and context of the terminology commonly used in this dissertation that may not be familiar to general reader, that was coined specifically in writing this work, or whose precise definition may vary depending on convention.

Since the raw signal is a voltage drop across a resistor, the unnormalized LIF profile out of the analysis $\varphi_{LIF}^{raw}(v[i], t[n])$ actually has units of volts and is often on the order of millivolts or tens of millivolts. Note that it is a function of the discrete set of values probed for velocity $v[i]$

and time $t[n]$, where i is the index of the velocities probed and n is the index of points in time. This signal is proportional to the intensity of LIF light by some unknown factor that depends on the details of the optics and electronics. The signal that is presented is normalized, meaning that the raw LIF profile $\phi_{LIF}^{raw}(v[i], t[n])$ is divided by a normalization factor $A(t[n])$ (in general, the normalization factor can change as a function of time):

$$\phi_{LIF}(v[i], t[n]) = \phi_{LIF}^{raw}(v[i], t[n]) / A(t[n]). \quad (4.1)$$

The term "**average signal normalization**" denotes that the LIF profile is normalized at all times by the average signal at the peak of the distribution. The normalizing factor is:

$$A(t[n]) = A = \max(\overline{\phi_{LIF}^{raw}(v[i], t[n])}), \quad (4.2)$$

where the bar indicates an average over time that is done independently for each velocity (i.e. the result is the time-averaged LIF profile). The max function is then evaluated over the velocity dimension to return a scalar value (the peak value of the time-averaged LIF profile). After this normalization, the unit of the normalized LIF profile ϕ_{LIF} can be considered dimensionless or arbitrary units (arb. u.). Since the signal is normalized by the same factor at all times, this normalization is useful to compare the signal level at one time to another. This normalization is the primary one used throughout this dissertation unless otherwise specified. It so happens in the cathode experiments of Chapter 6 that change in the raw signal level is much more apparent than change in the VDF itself. It also useful because the normalization is equivalent to the usual normalization of time-averaged LIF signal with a value of 1 at the peak of the distribution, enabling a simple comparison with the LIF profile from a lock-in amplifier.

On the other hand, "**integral normalization**" is used when the actual ion VDF is important and we wish to deemphasize the changes in signal level that may occur due to a change in ion density or metastable population. In this case, the normalization factor changes at each time step but is constant as a function of velocity to ensure that the integral over the normalized distribution is unity:

$$A(t[n]) = \int_{v_{min}}^{v_{max}} \phi_{LIF}^{raw}(v[i], t[n]) dv, \quad (4.3)$$

where v_{min} and v_{max} are the minimum and maximum velocity groups interrogated. Note that the integral is calculated numerically with the trapezoid rule because the LIF profile is measured at a discrete set of points in velocity. Though this normalization yields a proper distribution function

and emphasizes changes in the distribution rather than changes in signal level, the signal level at two different times cannot be directly compared since the normalization factor varies. As a proper velocity distribution (which is a probability density function), in this case the unit of the normalized LIF profile ϕ_{LIF} is s/m since it represents the probability density per unit velocity that a particle has a velocity between v and $v + dv$.

“**Chunk**” is the term commonly used for the shorter time-series over which averaging is performed, typically about 1 ms. The complete dataset for a single laser wavelength streamed for 60 s contains tens of thousands of chunks that are used to calculate either an average transfer function estimator or the triggered average LIF signal.

The term "**TRLIF**" is an abbreviation for "**time-resolved laser-induced fluorescence.**" "**TRLIF signal**" or "**TRLIF light**" refer to the actual fluorescence light signal that may be measured by a number of different techniques. It is the ideal signal that all TRLIF techniques seek to measure. The phrase "**TRLIF technique**" is an umbrella term used to refer to any technique capable of measuring the TRLIF signal, including all of the techniques discussed in Section 3.3.

"**TFLIF**" is short for "**transfer function laser-induced fluorescence.**" The term "**TFLIF technique**" is used to specifically refer to the new TRLIF technique that is the subject of this dissertation. Likewise, "**TFLIF system**" refers to the associated equipment and software implementing the technique at PEPL, and "**TFLIF signal**" refers to the signal measured using this system.

"**TALIF**" abbreviates "**triggered average laser-induced fluorescence.**" Similarly to the usage of TFLIF, "**TALIF technique**" refers to the combination of filtering, phase-sensitive detection, and triggered averaging presented as an alternative to the TFLIF technique in this chapter. "**TALIF signal**" refers to the signal measured using this technique. Usage of "**TALIF system**" will be rare because it is designed to use the same hardware as the TFLIF system; the only difference between the two systems is the analysis software. Note that the abbreviations "TALIF" and "TA-LIF" have been used in related literature by other authors to refer to "two-photon absorption LIF" [88] or "time-averaged LIF" [82], but those abbreviations are not related to the usage of "TALIF" in this dissertation.

4.2 Apparatus

A diagram of the experimental setup is shown in Figure 4.1. The laser probes the $5d^2F_{7/2} - 6p^2D_{5/2}^0$ transition of the singly-charged xenon ion (Xe II) at a wavelength of 834.7 nm (in air). The corresponding metastable Xe II ion velocity distribution function (VDF) is measured by collecting fluorescence from the $6s^2P_{3/2} - 6p^2D_{5/2}^0$ transition at 541.9 nm. We use a CW tunable diode laser (Toptica TA Pro) that has a typical output power of about 200 mW near the working wavelength range and a 20-50 GHz mode hop free range.

The beam is sampled at a few locations and the sample beams are sent to various diagnostics. (1) A Burleigh SA-91 etalon assembly with 2-GHz free spectral range ensures single-mode laser operation. (2) A HighFinesse WS/7 wavemeter with an accuracy of 60 MHz measures the wavelength. (3) An opto-galvanic cell (Hamamatsu L2783-42 XeNe-Mo galvatron) provides a stationary wavelength reference. (4) A Thorlabs PDA36A photodiode before the AOM monitors laser power.

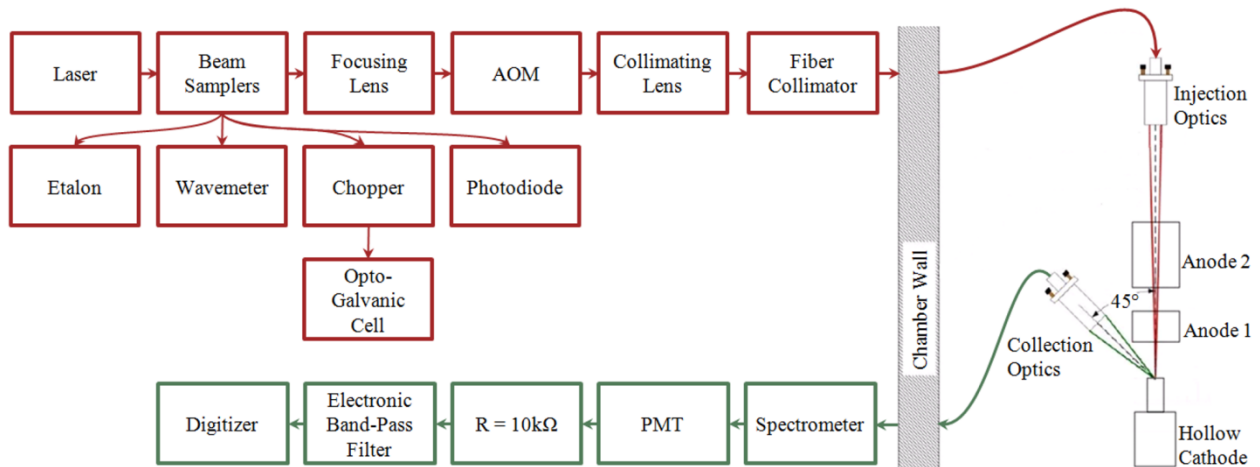


Figure 4.1. Diagram of the experimental setup showing a block diagram of airside optics sans mirrors, a schematic of the optics and hollow cathode plasma source inside the chamber, and a block diagram of instruments post collection outside the chamber.

The main laser beam is modulated by a NEOS 23080-1 acousto-optic modulator (AOM) that permits laser modulation frequencies up to about 5 MHz without significant distortion of the modulation waveform. A Thorlabs PDA10A photodiode after the AOM monitors the modulation waveform. A pair of 200-mm-focal-length focusing and collimating lenses provide a reasonable balance between diffraction efficiency ($\sim 70\%$) and rise time (32 ns) for the AOM. Following the

AOM, the beam is coupled to a 50- μm optical fiber with a numerical aperture of 0.22 and delivered into the chamber by an optical fiber feed through.

The beam is injected axially into the plasma source and focused down to a beam waist of approximately 1 mm in diameter at the interrogation volume. A 75-mm-diameter lens with 85-mm focal length images light collected from the interrogation volume onto a 1-mm optical fiber with a unity magnification. The interrogation volume, defined by the intersection of laser injection and light collection, is approximately cylindrical with a diameter of 1 mm and a length of 1 mm.

Collected light in the fiber is fed out of the chamber and into a SPEX-500M spectrometer, set to pass wavelengths near the 541.9 nm LIF transition with a bandwidth of about 1 nm. Light is then converted to an electrical signal by a Hamamatsu R928 photomultiplier tube (PMT), which is simply terminated by a 10-k Ω resistor for fast response. To remove noise away from the modulation frequency, a Krohn-Hite 3945 electronic filter provides fourth-order Butterworth band-pass filtering with cutoff frequencies 10% above and below the modulation frequency. Finally, the output from the filter is digitized by an Alazartech 9462 digitizer set to stream continuously to disk for 60 s per wavelength. Input and output amplification on the filter are set to fill the digitizer input range as well as possible without excessive saturation. The digitizer has 16-bit resolution and a maximum sample speed of 180 MHz, and a selection of input full scale ranges from ± 200 mV to ± 4 V. The discharge current, measured by a Tektronix TCP 312 current probe, is simultaneously sampled by the second channel of the 9462 digitizer for use in post-processing.

Since the data rate at the full 180 MHz sample speed is nearly 1 GiB/s, we solve the considerable data transfer and storage requirements by streaming directly to an array of 10 hard drives in RAID 6 with a net capacity of 16 TB, or 14.6 TiB. After optimizing system parameters, we found good results with a 30-MHz sample rate; thus, each 60-s measurement results in a 6.7 GiB data set in two-channel operation. A custom-built PC with dual hexacore Intel Xeon CPUs and 72 GiB of RAM houses the digitizer and RAID for data acquisition and analysis. The specifications enable it to analyze multiple data sets in parallel to save time.

4.3 Signal Processing

A simplified block diagram of the main steps in signal processing to recover the TRLIF signal from the noise is shown in Figure 4.2. The new method uses high speed laser modulation

(on the order of megahertz, well above the oscillation frequencies of interest), allowing two stages of as signal conditioning before the main averaging scheme. First, band-pass filtering improves SNR before digitization. After digitization, phase-sensitive detection (with a short time constant) demodulates the signal and provides a small boost to the SNR. There are two branches of signal processing following phase-sensitive detection. First, the demodulated signal with improved but still extremely poor SNR is used as the "output signal" together with an "input signal" such as the discharge current to calculate transfer function estimators. Unlike the techniques described in the previous chapter, which generally assume a periodic signal, the assumption here is that the input and output are related by a linear system (see Subsubsection 4.3.3.1). The series of transfer function estimators are ultimately averaged together to improve the SNR. A characteristic LIF output signal can then be synthesized for any input signal using a high-SNR average transfer function estimator. Alternatively to transfer function averaging, triggered averaging of the signal waveform in the time domain is also used to compare with the transfer function result as part of the validation test.

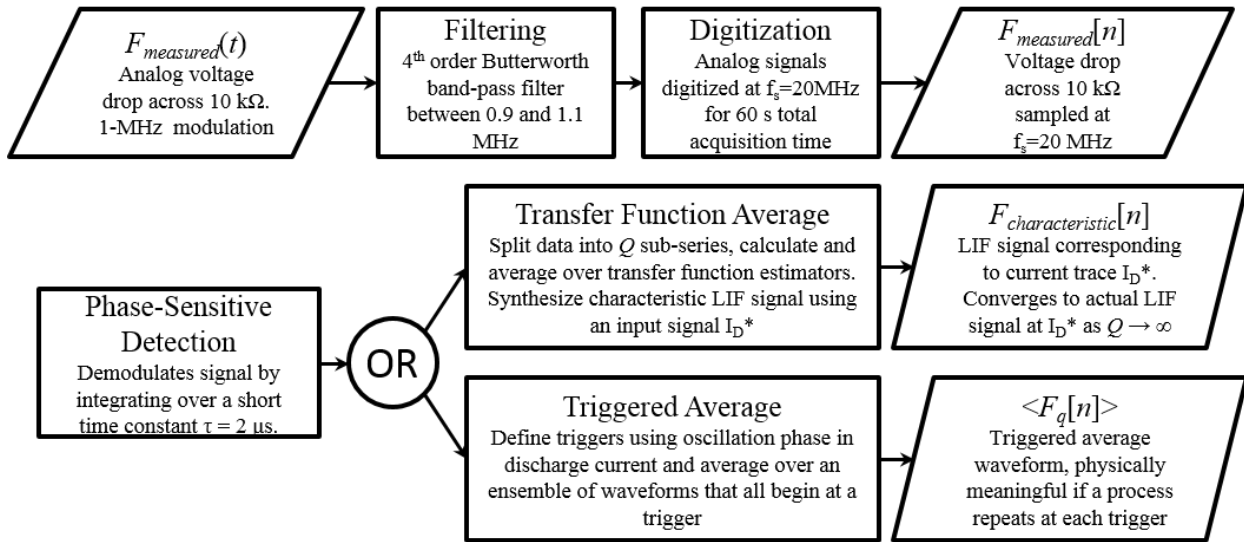


Figure 4.2. Flowchart of the post-processing steps on the PMT signal to recover the TRLIF signal. Though not explicitly shown in the flowchart, note that a high-speed signal related to the LIF signal (e.g. discharge current or floating probe) is necessary either to form transfer function estimators or to define the triggers used for triggered averaging.

It will be helpful to keep in mind a model of the voltage measured at the PMT terminating resistor as:

$$V(t) = F(t)M(t) + V_N(t), \quad (4.4)$$

where $F(t)$ is the envelope of the TRLIF signal (i.e. the signal that would exist for a CW laser at constant power), $M(t)$ is the laser modulation waveform, and $V_N(t)$ is some additive noise signal with a much larger RMS than the TRLIF signal. The laser modulation waveform may be an arbitrary function that is bounded between zero and one and periodic at the modulation frequency f_{ref} , but typically is a sine or square wave. The noise $V_N(t)$, or background signal, comes primarily from non-LIF light emitted in the plasma that the optics collect in addition to the LIF light. As mentioned previously, the noise light is orders of magnitude more intense than LIF (see Subsection 5.5.2), necessitating a signal recovery technique. The noise is random but has a well-defined spectral density generally weighted at frequencies below 1 MHz (for example, Figure 5.11 shows noise spectral densities measured from the H6).

In the following development of the signal processing theory, we calculate the *SNR* (signal-to-noise ratio), *I* (SNR improvement factor), and *ENBW* (equivalent noise bandwidth) at a number of different steps in the signal processing. The signal processing step with which the *SNR*, *I*, or *ENBW* are associated are given by the subscripts "raw" (raw signal measured at the PMT terminating resistor), "BP" (band-pass filter), "PSD" (phase-sensitive detection), "TF" (transfer function average), and "TA" (triggered average). For example, SNR_{BP} is the signal-to-noise ratio following the band-pass filter.

The goal of each of the following subsections is to describe the process that occurs at each step and the significance of it in the analysis, as well as to estimate the SNR improvement factor at each step. We model the final output SNR as a series of improvement factors for each step multiplied by the initial SNR in the raw PMT signal. The first three subsections cover the band-pass filter, phase-sensitive detection, and the Fourier analysis underpinning the transfer function average. Subsection 4.3.4 analyses the triggered average technique used as an alternative to the transfer function average. Subsection 4.3.5 proposes a benchmark based on a theoretical property linking the transfer function average to the triggered average even when the triggered average is not physically meaningful due to a nonperiodic signal.

4.3.1 Band-Pass Electronic Filter

A band-pass filter with a passband centered on the laser modulation frequency passes the TRLIF signal (first term in Equation (4.4)) while greatly reducing the frequency components of the noise (the second term) outside the passband. It improves the SNR of the digitized signal, thereby allowing a better effective voltage resolution. In other words, if the RMS noise is

reduced then the digitizer can use a smaller voltage input range, which reduces the size of digital voltage step compared to the signal scale.

The band-pass filter used for this signal conditioning, provided by the Krohn-Hite 3945, is composed of 4th-order Butterworth high-pass and low-pass filters in series. The goal of this subsection is to provide a theoretical basis for the function of the band-pass filter in the TFLIF system and to estimate appropriate values for the bandwidth. The filter bandwidth is the difference between the cutoff frequencies of the high-pass and low-pass filters (note that these are the -3dB points of the individual filters, which are slightly different from the -3 dB points of the series band-pass filter). We will consider filters with a bandwidth centered on the modulation frequency so that the cutoff frequencies are given by $f_{\pm} = f_{mod} \pm \Delta f/2$, where Δf is the bandwidth.

The transfer function $H(i\omega)$ of a filter is a complex number that is a function of frequency. The magnitude gives the gain of the filter and the argument gives the phase shift. The amplitude gain profile for the n^{th} -order series band-pass filter described above is given by:

$$|H(i\omega)|^2 = \frac{1}{1 + \left(\frac{\omega}{\omega_-}\right)^{2n}} \frac{1}{1 + \left(\frac{\omega_+}{\omega}\right)^{2n}}. \quad (4.5)$$

To give a sense of the overall frequency response of the filter, a Bode plot for the band-pass filter in a broad range of frequencies for $f_{mod} = 1$ MHz and $\Delta f = 100$ kHz is plotted in Figure 4.3. The filter has a linear response in the sense that an input signal at a constant frequency is mapped to an output signal at the same frequency with amplitude and phase offset shifted by the appropriate values in Figure 4.3. A more complicated input signal with some combination of Fourier frequency components will have each frequency component independently transformed by the appropriate gain and phase shift.

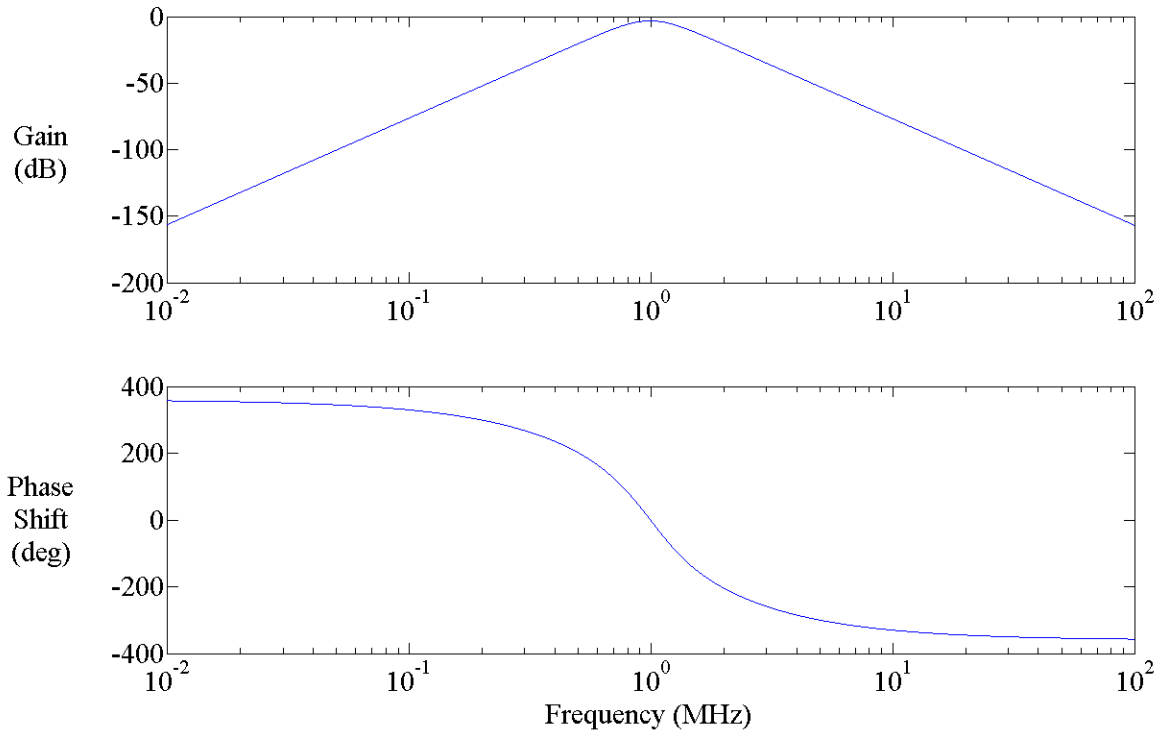


Figure 4.3. Bode plot of the frequency response of the series high-pass/low-pass filter with cutoff frequencies $f_{mod} \pm 0.1 f_{mod}$ for $f_{mod} = 1$ MHz.

The filter with $\Delta f = 100$ kHz was ultimately chosen as an appropriate filter for the measurements of this dissertation at a modulation frequency of $f_{mod} = 1$ MHz. Subsubsection 4.3.1.1 derives a formula that can be used to calculate the SNR improvement factor for a given filter and noise spectral density, while Subsubsection 4.3.1.2 discusses the equivalent noise bandwidth, which is a convenient approximation to compare the noise rejection of different filters. Subsubsection 4.3.1.3 calculates the SNR improvement factor for different filters as a function of bandwidth and finds that it is a weak function of bandwidth and there is no local maximum in SNR improvement. Subsubsection 4.3.1.4 calculates the bandwidth of the modulated signal that is necessary to pass without distortion. The result leads Subsubsection 4.3.1.5 demonstrating that distortion is negligible within the required band for the filter with $\Delta f = 100$ kHz at $f_{mod} = 1$ MHz. The $\Delta f = 100$ kHz filter was a reasonable choice because SNR improvement would not significantly increase for a narrower bandwidth, while a narrower bandwidth could potentially give borderline levels of distortion of the signal envelope.

4.3.1.1 Signal-to-Noise Ratio Improvement Factor

To calculate the SNR improvement factor, first consider the RMS LIF signal level following filtering. The RMS signal following the filter, $V_{s,BP}$, is:

$$V_{s,BP} = G |H_{BP}(i\omega_{mod})| V_{s,raw}, \quad (4.6)$$

where G is the collective gain of input/output amplifiers of the Krohn-Hite 3945, $|H_{BP}(i\omega_{mod})|$ is the gain of the band-pass filter itself at the modulation frequency (or the magnitude of the transfer function evaluated at the modulation frequency), and $V_{s,raw}$ is the RMS TRLIF signal voltage at the terminating resistor.

Second, the RMS noise levels before and after band-pass filtering can be calculated by integrating over the spectral noise density and gains. We summarize both equations with a subscript X that can be either "BP" to indicate signal following filtering, or "raw" to indicate signal before filtering:

$$V_{n,X} = \frac{1}{2\pi} \sqrt{\int_0^{\infty} V_D^2(\omega) G_X^2 |H_X(i\omega)|^2 d\omega}, \quad (4.7)$$

where $V_D(f)$ is the noise spectral density in the raw resistor signal in units of V/\sqrt{Hz} . $G_{BP} = G$ and $G_{raw} = 1$ since there is no amplification without the filter. $|H_{BP}(i\omega)|$ is the magnitude of the band-pass filter transfer function and $|H_{raw}(i\omega)|$ is the magnitude of a low pass filter transfer function that would be used at the Nyquist frequency for the anti-aliasing in the absence of the band-pass filter.

As is a common convention with LIF signals, the SNR here is defined as the ratio of RMS signal to RMS noise. The definitions above can be used to calculate the SNR before and after the band-pass filter. The SNR improvement factor of the band-pass filter is then the ratio of the SNR before and after band-pass filtering:

$$\begin{aligned}
I_{BP} &= \frac{SNR_{BP}}{SNR_{raw}} = \frac{\frac{V_{s,BP}}{V_{n,BP}}}{\frac{V_{s,raw}}{V_{n,raw}}} \\
&= \frac{G |H_{BP}(i\omega_{mod})| V_{s,raw}}{\frac{G}{2\pi} \sqrt{\int_0^\infty V_D^2(f) |H_{BP}(i\omega)|^2 df}} \frac{\frac{1}{2\pi} \sqrt{\int_0^\infty V_D^2(\omega) |H_{raw}(i\omega)|^2 d\omega}}{|H_{raw}(i\omega_{mod})| V_{s,raw}} \\
&= \frac{|H_{BP}(i\omega_{mod})| \sqrt{\int_0^\infty V_D^2(\omega) |H_{raw}(i\omega)|^2 df}}{|H_{raw}(i\omega_{mod})| \sqrt{\int_0^\infty V_D^2(\omega) |H_{BP}(i\omega)|^2 d\omega}}.
\end{aligned} \tag{4.8}$$

The improvement factor is invariant to the magnitude of signal and noise but is highly sensitive to the noise spectral density profile. Note that this signal conditioning gives a large improvement when the noise spectral density is relatively low near the modulation frequency and high for some other frequency band. It was estimated that the improvement factor is about an order of magnitude larger for the actual noise spectral density than for white noise [89].

4.3.1.2 Equivalent Noise Bandwidth

The above description is exact in that it takes into account the noise spectral density profile as a function of frequency. Equivalent noise bandwidth (ENBW) is a convenient concept when the noise spectral density is approximately white (constant as a function of frequency) in the frequency band of interest.

In general, the equivalent noise bandwidth of a filter is defined as:

$$ENBW = \frac{1}{2\pi} \int_0^\infty \frac{|H(i\omega)|^2}{\max(|H(i\omega)|^2)} d\omega, \tag{4.9}$$

where, as above, $|H(i\omega)|^2$ is the square magnitude of the filter transfer function, which is the frequency dependent power gain of the filter. The ENBW of a filter is typically a small factor larger than the bandwidth defined by the cutoff frequencies of a band-pass or low-pass filter since the finite filter roll-off passes “extra” noise that would be cutoff for a brick-wall filter at the cutoff frequency. Tabulated values are available for common filters, but the series of low-pass

and high-pass filters used here are uncommon and the factor by which the ENBW differs from the Δf bandwidth varies if Δf is small enough that the transition regions overlap. If the Noise Spectral density is constant, then it can be factored out of the integral and the RMS noise calculation simplifies to:

$$V_n = V_D \max(|H(i\omega)|) \sqrt{ENBW}. \quad (4.10)$$

If also the gain at the modulation frequency is at or near the maximum in the filter gain (as it should be), then Equation 3.5 simplifies to:

$$I = \frac{\sqrt{ENBW}}{\sqrt{ENBW_{raw}}}, \quad (4.11)$$

a well-known result for the SNR improvement of filtering.

Note that this formula is not useful for calculating the SNR improvement of the BP filter since the noise spectral density is not white (see Subsection 5.5.1). The concept of ENBW is still defined for a filter regardless of the actual noise spectral density, and its definition is included here because a need for $ENBW_{BP}$ arises in the improvement factor of phase-sensitive detection in Section 4.3.2. The use of ENBW does apply there because the noise spectral density is approximately white within the narrow bandwidth of the band-pass filter.

4.3.1.3 Optimal Bandwidth in Terms of Signal-to-Noise Ratio Improvement

The main parameter to consider for the band-pass filter is the bandwidth Δf of the pass band, defined as the difference between the upper and lower cutoff frequencies. An optimal setting would be one that maximizes the SNR improvement factor while the distortion of the signal is minimal.

Ignoring distortion of the time-resolved signal and considering a theoretical brick-wall filter, the answer is simple: an infinitesimally narrow pass band centered on f_{mod} will reject almost all broadband noise while fully passing the signal. The problem is more complicated for real-world filters that have some transition region and roll-off instead of a sharp transition at the cutoff frequencies. Figure 4.4 shows examples of gain profiles for bandpass filters implemented with 4th-order Butterworth high-pass and low-pass filters in series. As the bandwidth of the pass band is reduced and the upper and lower cutoff frequencies approach the modulation frequency ($f_{mod} = 1$ MHz for this example and most of the experiments in this dissertation), the modulation

frequency enters the transition region and the signal gain is significantly reduced from unity. In addition, the finite roll-off ensures a significant ENBW even for $\Delta f = 0$ Hz.

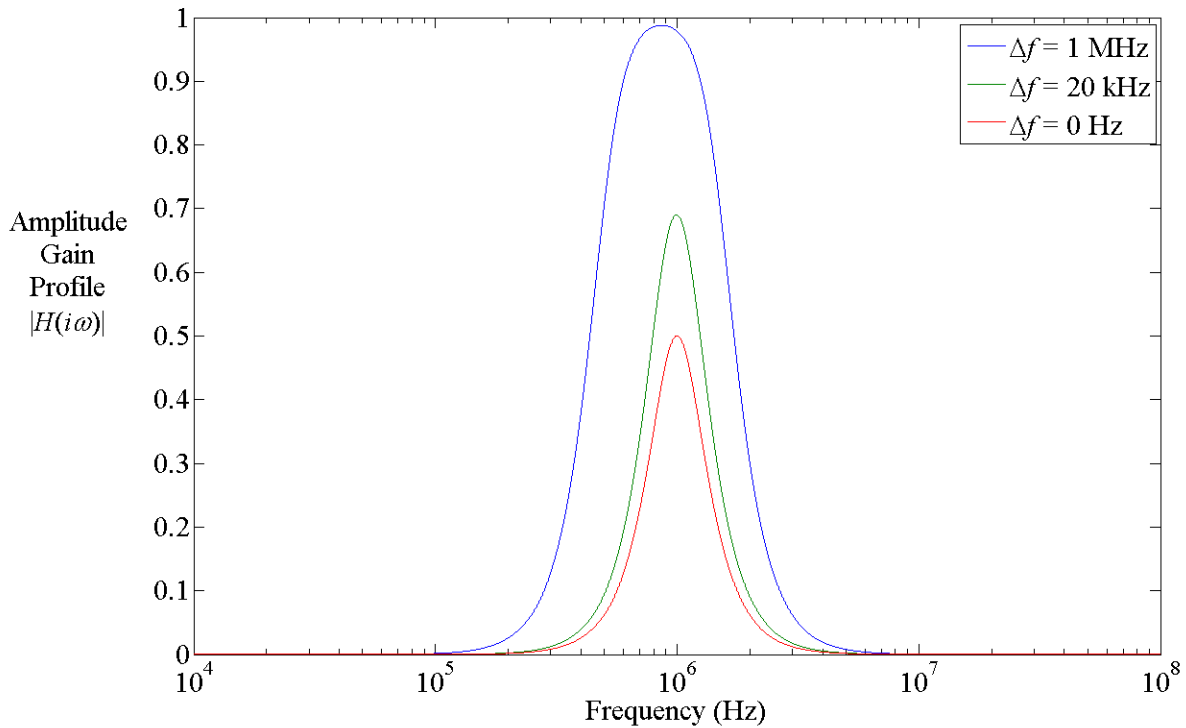


Figure 4.4. Amplitude gain profiles of band-pass filters centered on 1 MHz for three representative bandwidths from 0 to 1 MHz.

The first question is whether the SNR continues improving as the pass band is reduced to zero. In other words, does the improvement in noise rejection from a narrower pass band always outweigh the loss in signal from the decreasing signal gain at $f_{mod} = 1$ MHz, or does the loss of signal eventually outweigh the noise rejection, leading to a local maximum in I_{BP} versus Δf ? The answer depends on the gain profile of the filter and the noise spectral density in the region of the pass band as shown by Equation (4.8). For simplicity, suppose the noise spectral density is approximately white (constant density as a function of frequency) in the region of the band-pass filter, then the answer depends only on the gain profile of the filter. This is likely a good approximation for narrow filters that only span a short frequency range but possibly not a good approximation for a wide bandwidth.

The SNR improvement factor calculated by numerically solving Equation (4.8) for a range of bandwidths is shown in Figure 4.5. The code was validated by reproducing the known ENBW

of low-pass filters. The SNR improvement factor monotonically increases but converges to a value of about 5.5 as the bandwidth is reduced to zero. The result means that the loss of noise from reducing the ENBW of the filter nearly cancels the loss of signal from reduced peak gain for very narrow-bandwidth filters. The SNR improvement factor can be about an order of magnitude larger for the non-white noise spectral density of an actual plasma source [89], but it will still converge to some value since there is still a finite gain profile at $\Delta f = 0$.

Two important consequences follow from the result. First, there is theoretically no local maximum in the SNR improvement beyond which the SNR actually falls for a tighter bandwidth, and therefore an acceptable filter bandwidth Δf is limited only by the distortion within the desired signal bandwidth that a narrower filter introduces. Second, the improvement factor weakly depends on the bandwidth only weakly for reasonable bandwidths, with about a 15% difference between $\Delta f = 500$ kHz and $\Delta f = 0$ Hz.

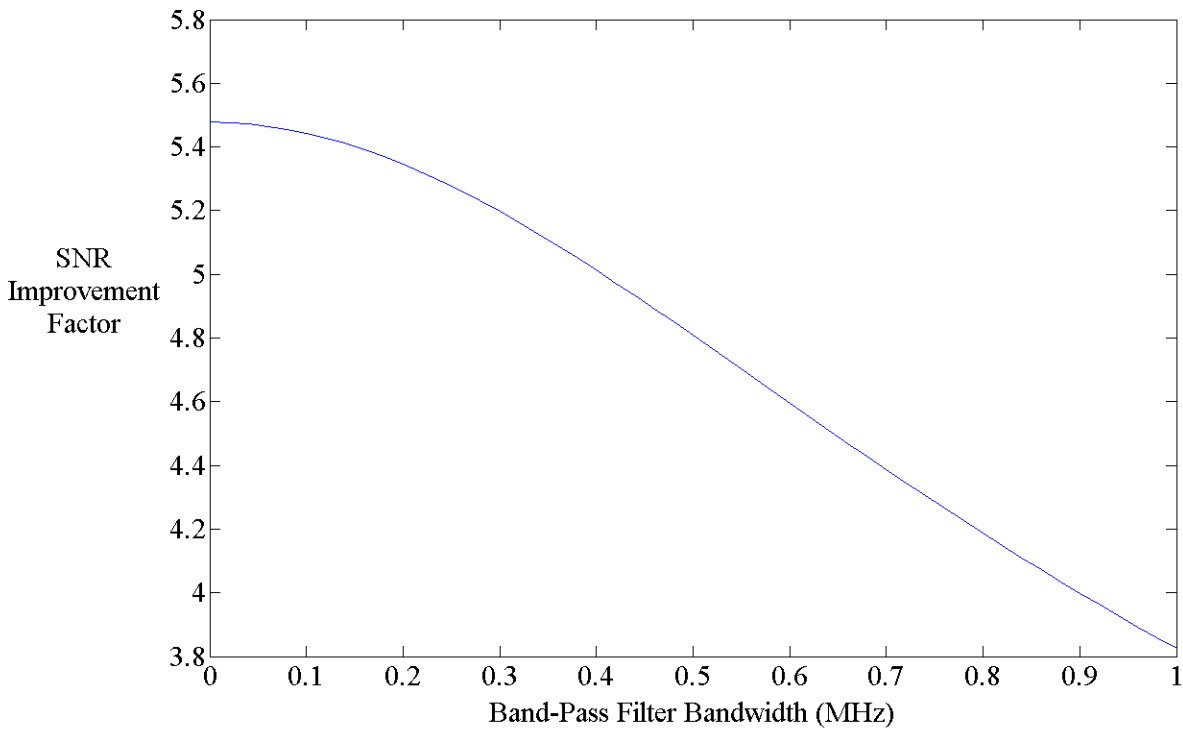


Figure 4.5. The SNR improvement factor calculated by solving Equation (4.8) for $f_{mod} = 1$ MHz and varying the bandwidth of a band-pass filter centered on f_{mod} . The calculation, assuming white noise, shows that the SNR improvement monotonically increases but converges to a value of about 5.5 as the bandwidth approaches zero.

Regardless of the SNR improvement from the band-pass filter, the final SNR of the TFLIF signal may not depend on the band-pass filter bandwidth because, as discussed in the following subsection, phase-sensitive detection will have a narrower transmission window that will reject more of the noise spectrum anyway. The final SNR could depend on the filter bandwidth only when the RMS noise is sufficiently larger than the RMS signal that the effective voltage resolution of the signal is small enough that quantization noise becomes significant. The main purpose of the band-pass filter, therefore, is to increase the dynamic range of the system by rejecting some portion of the noise before digitization and thereby raise the effective voltage resolution because a smaller voltage scale can be used to digitize the signal (see Subsubsection 5.1.1.2). As long as the system is out of the regime of quantization noise, however, an increase in the band-pass filter SNR improvement will not result in a corresponding improvement in the final SNR of the TFLIF signal.

4.3.1.4 Bandwidth of the Modulated Signal

A second question to address is at what point a narrow band-pass filter begins to distort the signal envelope that we seek to measure. The first step to answering this question is to define the bandwidth containing the modulated signal. Recall from Equation (4.4) that a model of the voltage measured at the PMT terminating resistor is:

$$V(t) = F(t)M(t) + N_{raw}(t), \quad (4.12)$$

where $F(t)$ is the envelope of the TRLIF signal (i.e. the signal that would exist for a CW laser at constant power), $M(t)$ is the laser modulation waveform, and $N_{raw}(t)$ is some additive noise signal with a much larger RMS than the TRLIF signal. Consider the Fourier transform of this signal and the convolution theorem:

$$\tilde{V}(f) = \tilde{F}(f) \otimes \tilde{M}(f) + \tilde{N}_{raw}(f). \quad (4.13)$$

In frequency space, the signal is the convolution (denoted by \otimes) of the Fourier transforms of the signal envelope $F(t)$ and the modulation waveform $M(t)$. Convolution is a mathematical operation that generally “mixes” two functions and is defined by:

$$f(t) \otimes g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau. \quad (4.14)$$

If the modulation is purely sinusoidal, then its frequency domain description is a pair of Dirac delta functions (recall that real-valued functions have symmetric positive and negative

frequency components in the complex exponential form of the Fourier transform). Let the modulation waveform be a cosine with frequency f_{mod} and phase offset φ_{mod} :

$$\begin{aligned} M(t) &= 1 + \cos(2\pi f_{mod}t + \varphi_{mod}) \\ \tilde{M}(f) &= \delta(f) + \frac{\delta(f - f_{mod})e^{i\varphi_{mod}} + \delta(f + f_{mod})e^{-i\varphi_{mod}}}{2}. \end{aligned} \quad (4.15)$$

In this case, the signal bandwidth is retained but shifted both left and right into two bands centered on $f = \pm f_{mod}$ instead of $f = 0$:

$$\begin{aligned} \tilde{V}_{raw}(f) &= \tilde{F}(f) \otimes \left(\delta(f) + \frac{\delta(f - f_{mod}) + \delta(f + f_{mod})}{2} \right) + \tilde{N}_{raw}(f) \\ &= \tilde{F}(f) + \frac{\tilde{F}(f - f_{mod}) + \tilde{F}(f + f_{mod})}{2} + \tilde{N}_{raw}(f) \\ &= \tilde{F}(f) + \frac{\overline{\tilde{F}(-f + f_{mod})} + \tilde{F}(f + f_{mod})}{2} + \tilde{N}_{raw}(f). \end{aligned} \quad (4.16)$$

Note that the final step explicitly showing complex conjugate symmetry about the y-axis holds if $\tilde{F}(f)$ itself is complex conjugate symmetric about the y-axis, which is true for any real-valued signal envelope function $F(t)$. The first term in Equation (4.16) is the Fourier transform of the envelope and is present at the input of the band-pass filter but is not passed because the passband is near the modulation frequency. The result, illustrated in Figure 4.6, is that if the signal envelope has a bandwidth within the interval $[-f_B, f_B]$, where f_B is the single-sided bandwidth of the signal, then the modulated signal will have a bandwidth in the interval $[f_{mod} - f_B, f_{mod} + f_B]$ and $[-f_{mod} - f_B, -f_{mod} + f_B]$. Therefore, the band-pass filter must pass the signal with minimal gain and phase distortion within that band.

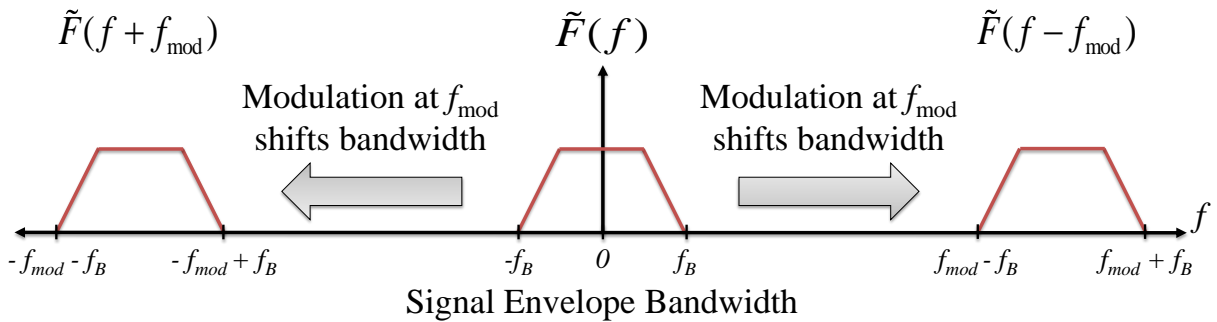


Figure 4.6. An illustration showing that signal modulation of the envelope with a cosine waveform shifts the signal bandwidth from centered on zero to two sidebands centered on the modulation frequency f_{mod} with the same bandwidth as the double-sided bandwidth of signal envelope $F(t)$.

4.3.1.5 Distortion from the Band-Pass Filter on the Modulated Signal

We have determined above the bandwidth of the modulated signal. To faithfully pass the modulated signal, there must be negligible distortion within this bandwidth. Now we determine the distortion that a given filter may introduce into a signal in that band. Consider the band-pass filter described earlier that is composed of 4th-order Butterworth high-pass and low-pass filters in series. For this example, we take $f_{mod} = 1$ MHz and the cutoff frequencies of the individual filters to be $f_{mod} \pm 100$ kHz, the modulation frequency and filter used for the experiments of Chapter 6 and Chapter 7.

The experiments in Chapter 6 have a signal contained within the single-sided bandwidth $f_B = 20$ kHz, hence we consider the filter's frequency response in the interval $[f_{mod} - 20$ kHz, $f_{mod} + 20$ kHz], shown in Figure 4.7. The goal is to show that the distortion of the modulated signal is negligible within this band.

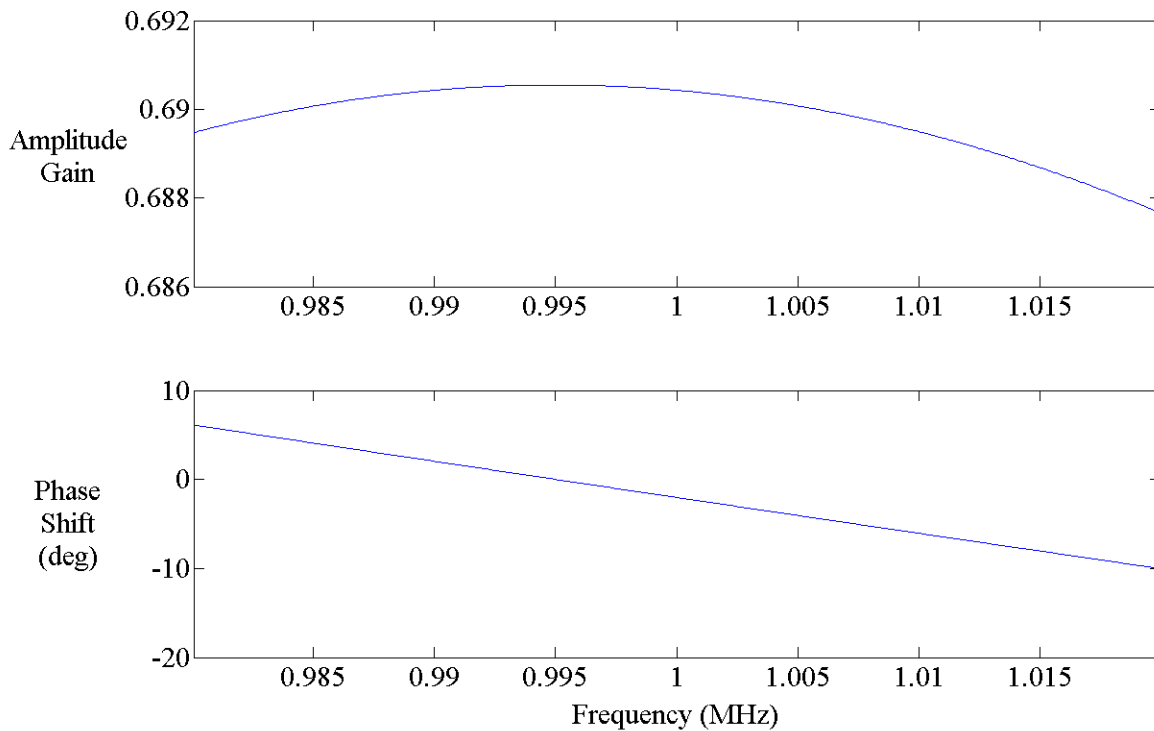


Figure 4.7. Gain and phase shift from the Bode plot of Figure 4.3, zoomed in on the bandwidth interval $[f_{mod} - 20$ kHz, $f_{mod} + 20$ kHz].

There are two types of distortion that may be introduced by the filter. First, distortion could occur if two frequency components are passed with a significant difference in gain. Figure

4.7 shows that this is not a problem in the bandwidth of interest since gain varies by less than 1%. Though the gain is significantly less than unity, it effectively applies the same scaling factor to all components and does not appreciably affect the envelope waveform.

Second, there may be distortion from variation in the phase shift between about -10° and 5° in the band of interest. The objective is to determine how this phase shift in individual frequency components may distort the resulting modulated envelope waveform. It is illustrative to consider the modulation of a signal envelope of a cosine at frequency f_{en} and phase offset φ_{en} :

$$\begin{aligned} F(t) &= A \cos(2\pi f_{en} t + \varphi_{en}) \\ \tilde{F}(f) &= A \frac{\delta(f - f_{en})e^{i\varphi_{en}} + \delta(f + f_{en})e^{-i\varphi_{en}}}{2}. \end{aligned} \quad (4.17)$$

The second term of Equation (4.16), the only part of the desired signal to be passed by the band-pass filter, becomes:

$$\begin{aligned} & \frac{\overline{\tilde{F}(-f + f_{mod})e^{i\varphi_{mod}}} + \tilde{F}(f + f_{mod})e^{-i\varphi_{mod}}}{2} \\ &= A \frac{\delta(-f + f_{mod} - f_{en})e^{i\varphi_{mod}}e^{-i\varphi_{en}} + \delta(-f + f_{mod} + f_{en})e^{i\varphi_{mod}}e^{i\varphi_{en}}}{4} \\ &+ A \frac{\delta(f + f_{mod} - f_{en})e^{-i\varphi_{mod}}e^{i\varphi_{en}} + \delta(f + f_{mod} + f_{en})e^{-i\varphi_{mod}}e^{-i\varphi_{en}}}{4}. \end{aligned} \quad (4.18)$$

As this signal is passed through the filter, each term is transformed by the appropriate gain and phase shift since the delta functions each correspond to a fixed frequency:

$$\begin{aligned} & \frac{\overline{\tilde{F}(-f + f_{mod})e^{i\varphi_{mod}}} + \tilde{F}(f + f_{mod})e^{-i\varphi_{mod}}}{2} \\ &= A \frac{G_- e^{i\phi_-} \delta(-f + f_{mod} - f_{en})e^{i\varphi_{mod}}e^{-i\varphi_{en}} + G_+ e^{i\phi_+} \delta(-f + f_{mod} + f_{en})e^{i\varphi_{mod}}e^{i\varphi_{en}}}{4} \\ &+ A \frac{G_- e^{-i\phi_-} \delta(f + f_{mod} - f_{en})e^{-i\varphi_{mod}}e^{i\varphi_{en}} + G_+ e^{-i\phi_+} \delta(f + f_{mod} + f_{en})e^{-i\varphi_{mod}}e^{-i\varphi_{en}}}{4}, \end{aligned} \quad (4.19)$$

where the transfer function gain and phase shift variables are:

$$H(i\omega) = \begin{cases} G_- e^{i\phi_-} & \text{for } f = f_{mod} - f_{en} \\ G_- e^{-i\phi_-} & \text{for } f = -(f_{mod} - f_{en}) \\ G_+ e^{i\phi_+} & \text{for } f = f_{mod} + f_{en} \\ G_+ e^{-i\phi_+} & \text{for } f = -(f_{mod} + f_{en}). \end{cases} \quad (4.20)$$

The numerical values of the gain and phase shift can be found in the Figure 4.7. We can recognize the Fourier transform of the sum of two cosines to transform Equation (4.20) back to the time domain. Assume the gain factors are equal, a very good approximation as noted above. Then we can show by applying a trigonometric identity:

$$V_{BP} \approx AG_+ \cos(2\pi f_{\text{mod}} t + \frac{\phi_- + \phi_+}{2} + \varphi_{\text{mod}}) \cos(2\pi f_{\text{en}} t + \frac{\phi_+ - \phi_-}{2} + \varphi_{\text{en}}) + N_{BP}(t). \quad (4.21)$$

The signal passed by the band-pass filter is essentially the same as the input except both the envelope and the carrier are shifted by a phase offset term. The phase offsets can be converted into time delays known as the group delay and phase delay:

$$\begin{aligned} V_{BP} &\approx AG_+ \cos(2\pi f_{\text{mod}}(t - \tau_\varphi) + \varphi_{\text{mod}}) \cos(2\pi f_{\text{en}}(t - \tau_g) + \varphi_{\text{en}}) \\ \tau_g &= -\frac{\phi_+ - \phi_-}{4\pi f_{\text{en}}} \approx 1.1 \mu\text{s} \\ \tau_\varphi &= -\frac{\phi_- + \phi_+}{4\pi f_{\text{mod}}} \approx 5 \text{ ns.} \end{aligned} \quad (4.22)$$

These are more general definitions for the group and phase delay of a modulated cosine than the typical textbook formulae. The textbook definitions of the delay times assume that the phase is a linear function of frequency, in which case the delay times are:

$$\begin{aligned} \tau_g &= -\frac{d\phi(\omega)}{d\omega} \\ \tau_\varphi &= -\frac{\phi(\omega)}{\omega}. \end{aligned} \quad (4.23)$$

Though the phase in this example is linear in the region of interest, as shown in Figure 4.7, the more general forms of the delays could be helpful in a more borderline case where the phase is not extremely linear, such as if we use the same filter to acquire a broader bandwidth.

One key conclusion of the above result is that both group and phase delay are approximately constant (independent of the envelope frequency) in a region where phase shift is approximately a linear function of frequency. This is an important result because the previous discussion holds not only for a modulated cosine but also for envelope functions composed of arbitrary frequency components in the approximately linear band. All of the frequency components in the envelope are delayed by the group delay time and each of them are all

modulated with the same phase delay so that there is no distortion of either the envelope or the carrier wave except for a slight shift in time.

The final task is to specify the actual value of the group and phase delay and determine whether they need to be taken into account. The group delay calculated by Equation (4.22) is about 1.1 μs . Numerical simulations of passing a modulated cosine through a digital filter corroborated this value. This is a small delay time compared with the scale of the ~ 10 kHz oscillations and is near the time resolution of the TFLIF system when used with the phase-sensitive detection parameters discussed in the next subsection, and therefore it was not taken into account in any results reported in this dissertation. There is no distortion of the TFLIF signal since all frequency components will be delayed by the same group delay, but the delay will desynchronize the TFLIF signal from other measurements, and therefore a comparison to discharge current, Langmuir probe, and FASTCAM data should ideally take the TFLIF group delay into account (as well as any possible delay in the other measurements). The phase delay was found to be about 5 ns from both Equation (4.22) and numerical simulations. This delay is negligible as it is even smaller than the digitizer's sampling period. Even a significant phase delay would not be a problem because the phase locking discussed in Section 5.2 locks to the actual measured signal phase without a reference to the original phase before the filtering.

4.3.2 Phase-Sensitive Detection

By the flowchart in Figure 4.2, the signal is digitized after the signal conditioning by the electronic band-pass filter. All signal processing following occurs digitally in post processing after the experiment. As discussed previously, the SNR is somewhat improved following band-pass filtering, but the TRLIF signal is still modulated at the laser modulation frequency and buried in noise. The desired TRLIF signal is the envelope $F(t)$ in Equation (4.4). Phase-sensitive detection (PSD) is applied in software to demodulate the signal so that the averaging procedure in the next analysis step can recover the envelope with high SNR.

Phase-sensitive detection is the algorithm used by lock-in amplifiers to recover a small signal at a known reference frequency from a strong noise level or background signal. As discussed in Section 3.1, the standard method of time-averaged LIF measurements involves modulating the laser on the order of kilohertz with a mechanical chopper and then recovering the LIF signal with a lock-in amplifier set to integrate over a long time constant of at least about 100

ms to improve SNR sufficiently. The long time constant averages over time-varying information; hence a short time constant must be used if PSD is to be used with a time-resolved technique.

Phase-sensitive detection involves two steps. It multiplies the input signal by a reference signal at a known reference frequency (a sine wave reference is used in this work) and then applies a low-pass filter. The low-pass filter effectively integrates over a few time constants in time. The time constant τ depends on the cutoff frequency of the low-pass filter as $\tau = 1/(2\pi f_c)$, where f_c is the cutoff frequency.

4.3.2.1 Frequency Domain Interpretation

To understand how this process demodulates a signal, consider the case where the input modulated signal and reference signals, $V_{mod}(t)$ and $V_{ref}(t)$, are both sinusoidal and described by the same equation:

$$V_x(t) = \sqrt{2}V_{x,RMS} \sin(\omega_x t + \varphi_x), \quad (4.24)$$

where the subscript x is "mod" for the input modulated signal or "ref" for reference. $V_{x,RMS}$ is the RMS amplitude, ω_x is the angular frequency, and φ_x is the phase offset. By a trigonometric identity, the product of the input and reference signals contains components at the sum and difference of the two frequencies:

$$\begin{aligned} V_{mod}(t)V_{ref}(t) \\ = V_{mod,RMS} V_{ref,RMS} \left[\cos((\omega_{mod} - \omega_{ref})t + \varphi_{mod} - \varphi_{ref}) - \cos((\omega_{mod} + \omega_{ref})t + \varphi_{mod} + \varphi_{ref}) \right]. \end{aligned} \quad (4.25)$$

The key to phase-sensitive detection is prior knowledge of the input signal frequency to be recovered so that the reference can be set to the same frequency. Since $\omega_{mod} = \omega_{ref}$, there are frequency components at $2\omega_{mod}$ and DC. The low-pass filter cutoff frequency is set low enough to reject the $2\omega_{mod}$ component while passing the DC component. The net effect is that original sine wave input is converted to a DC signal proportional to the amplitude of the input, *i.e.* it is demodulated.

It is important to note that the output DC amplitude following the low-pass filter, V_{out} , is proportional to the cosine of the difference between the two phase offsets:

$$V_{out} = V_{mod,RMS} V_{ref,RMS} \cos(\varphi_{mod} - \varphi_{ref}). \quad (4.26)$$

Consequently, not only must the modulation frequency be known to high precision, but the reference phase should also be equal to the signal phase, hence the name "phase-sensitive detection." This requirement leads to the practical problem of phase locking when the signal is noisy, covered in Subsection 5.2. Fortunately, the cosine function is somewhat forgiving; the signal is above 98% of the maximum even at a difference of $\pm 10^\circ$. Of course, it falls rapidly after that point, necessitating a reasonably accurate measurement of the signal phase.

If we consider the input signal as composed of the signal to be recovered $V_{mod}(t)$ plus some noise with some distribution of spectral density, then the product will still have a DC component proportional only to the input sine wave amplitude $V_{mod}(t)$ as in Equation (4.26). This is true because the DC signal only appears for signal frequencies at the reference frequency. Even random incoherent noise near the reference frequency produces a zero mean output from Equation (4.26) because the phase of the noise randomly varies. A coherent interference signal near the reference frequency could produce a "false" DC output, so it is important to use a modulation frequency away from any sources of interference. This is also what makes resonant LIF tricky: scattered laser light can be registered by the lock-in amplifier just like LIF.

The low-pass filter acts to pass the desired DC signal while removing contributions from noise at other frequencies. We can think of a phase-sensitive detector as having a "transmission window" [90] that is centered on the signal frequency and that will pass the signal and noise components near the signal frequency that satisfy $|f_{mod} - f_{ref}| < f_c$. A lower filter cutoff frequency rejects more noise while passing the signal (thereby raising SNR). As we are about to see, however, this SNR improvement comes with a tradeoff from the time domain interpretation: reducing the cutoff frequency also raises the integration time constant, thereby destroying time-resolved information.

4.3.2.2 Time Domain Interpretation

A first order low-pass filter can be modeled by the following differential equation in the time domain:

$$\frac{dV_{out}}{dt} + \frac{V_{out}}{\tau} = \frac{V_{in}}{\tau}, \quad (4.27)$$

where V_{in} and V_{out} are the input and output voltages and $\tau = 1/(2\pi f_c)$ is the time constant. For an arbitrary continuous input voltage function with an initial value of $V_{out}(t = 0) = 0$, the particular solution is:

$$V_{out}(t) = \frac{e^{-t/\tau}}{\tau} \int_0^t e^{t'/\tau} V_{in}(t') dt' . \quad (4.28)$$

This is a weighted average over the input voltage with a weight function $w(t) = \exp(t/\tau)$, in other words an exponentially weighted average. This is a more precise statement that phase-sensitive detection (with a low-pass filter) averages the input over a few τ in time. Instead of using a low-pass filter, which leads to the exponentially weighted average, the averaging could alternatively be done with a uniformly weighted average integral or a direct average of the mixed input signal and reference over a better defined length of time. The low-pass filter is used here since it is the classic method of phase-sensitive detection, easy to implement in MATLAB, not computationally intensive, and easy to analyze theoretically since it has a well-defined gain profile.

Theoretically, the exponentially weighted average is over all time, but the weight factor at $t' = t - 5\tau$ is less than 1% of the weight factor at $t' = t$, hence the average is practically carried out over four or five time constants, with earlier input signal giving a negligible contribution unless the signal was much larger at earlier times. The settling time for a step change in the signal output with a first order filter is above 99% signal after 5τ time or 90% after 2.5τ time. A non-step change in the signal will be followed even more closely. Thus, the time resolution of the output is considered to be limited to about a few time constants, though the raw data are sampled much faster.

The preceding paragraph considers the input and output of a low-pass filter only. Consider the full phase-sensitive detector with the input voltage multiplied by a reference signal. In addition, consider that the input voltage can be decomposed into a form of the Fourier transform with real-valued sine and cosine components at all frequencies:

$$V_{in}(t) = \int_0^\infty a(f) \cos(2\pi ft) + b(f) \sin(2\pi ft) df . \quad (4.29)$$

Including the reference sine wave factor and the Fourier decomposition of the input signal, the exponentially weighted average results in:

$$V_{out}(t) = \frac{e^{-t/\tau}}{\tau} \int_0^t e^{t'/\tau} \sin(\omega_{ref} t') \int_0^\infty a(f) \cos(2\pi f t') + b(f) \sin(2\pi f t') df dt'. \quad (4.30)$$

This form is highly suggestive of the orthogonality of sine and cosine functions. By the orthogonality of sines and cosines, the integral over t' will tend to zero for all Fourier frequency components except those near the reference frequency. Hence, phase-sensitive detection picks out the component in the input signal at the reference frequency by exploiting the orthogonality of sinusoidal function. Further, this form gives insight into how phase-sensitive detection is in fact phase-sensitive. Since sines and cosines are orthogonal, the out-of-phase component of the input at the reference frequency (the cosine term) is not recovered. This leads again to the conclusion that the phase of the signal must be known in addition to the frequency to eliminate the out of phase component of the signal. As a final note, the definition of orthogonality for functions is an integral from negative infinity to positive infinity. Then the integral of the exponentially weighted average will be a better approximation of that when the time constant is large enough that the averaging is done over many reference periods.

Now suppose that the input signal is a sine wave whose phase offset has been measured and the reference signal is a sine wave set to the same phase offset and frequency, which is the condition to maximize the output signal from Equation (4.26). Then Equation (4.28) becomes:

$$V_{out}(t) = \frac{e^{-t/\tau}}{\tau} \int_0^t e^{t'/\tau} \sin^2(\omega_{ref} t') dt'. \quad (4.31)$$

This can be solved analytically, or alternatively the low-pass filtering of the mixed input signal and reference can be done numerically in MATLAB to generate a solution equivalent to the analytic one. The solution for a few values of τ with a 1-MHz modulation frequency are plotted in Figure 4.8. When viewed from the time domain, the $2\omega_{mod}$ ripple from the frequency domain discussion can be interpreted as coming from the weight factor favoring different parts of the $\sin^2(\omega_{ref} t)$ waveform at different times. This form intuitively shows the need to use a time constant larger than the modulation period so that the weighted average integrates over many reference periods to reliably detect the DC signal and reject the transient part of the waveform. Most importantly, it also demonstrates the loss of time resolution from increasing the time constant since increasing the time constant increases the response time to a change in the signal envelope.

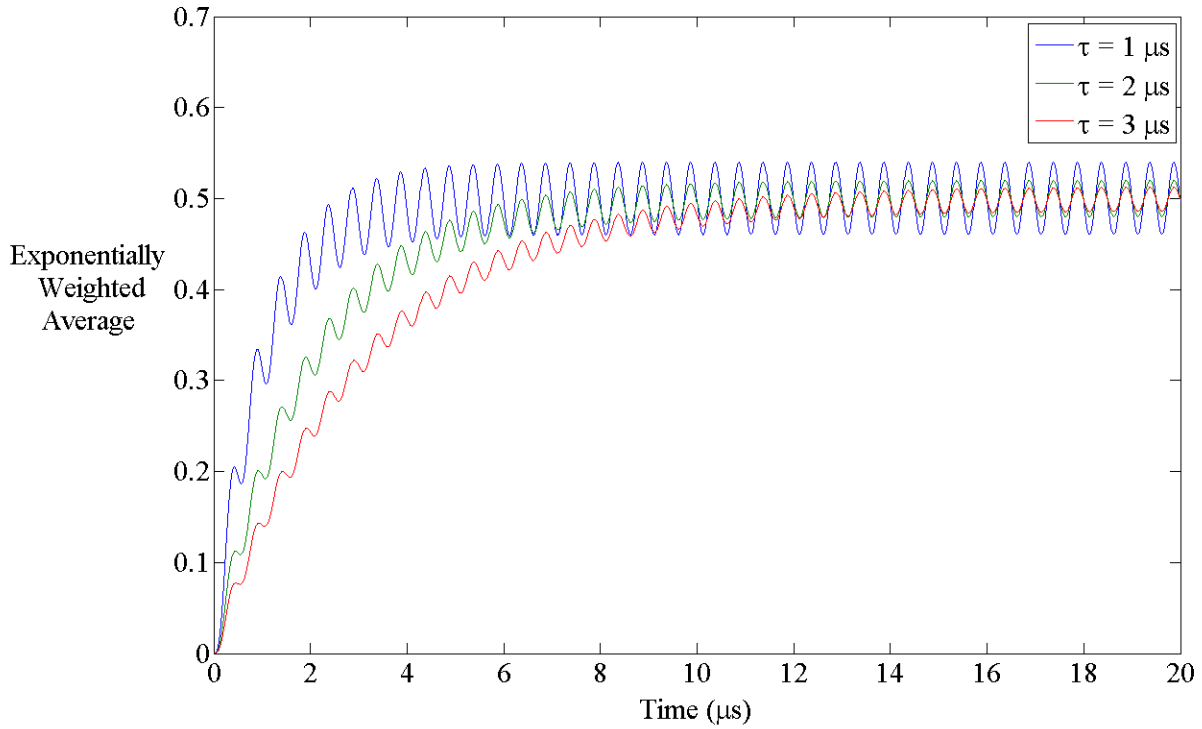


Figure 4.8. The output of phase-sensitive detection $V_{out}(t)$ given by solutions to Equation (4.31) for a few τ values, which show the increasing response time and decreasing ripple as τ increases.

4.3.2.3 Relevant Features for TRLIF Measurements

If the noise spectral density within the pass band of the band-pass filter is approximately white, which it is for a small pass band, then the SNR improvement for PSD (in terms of RMS voltages) can be shown to be the ratio of the equivalent noise bandwidths of the band-pass filter and the phase-sensitive detector [90]:

$$I_{PSD} = \frac{SNR_{PSD}}{SNR_{BP}} = \frac{\sqrt{ENBW_{BP}}}{\sqrt{ENBW_{PSD}}}. \quad (4.32)$$

The SNR improvement comes from reducing the width of the transmission window that passes noise near the reference frequency. Now, $ENBW_{PSD}$ is the bandwidth of the low-pass filter used in phase-sensitive detection. For a first-order filter, this is:

$$ENBW_{PSD} = \frac{\pi}{2} f_c = \frac{\pi}{2} \frac{1}{2\pi\tau} = \frac{1}{4\tau}. \quad (4.33)$$

Note that the improvement factor is proportional to the square root of the time constant.

Unlike the conceptual discussion above, the TRLIF signal is not a simple sine wave but a carrier wave at the reference frequency that is amplitude modulated within an envelope $F(t)$ as in Equation (4.4). If the envelope $F(t)$ is approximately constant on the integration time scale of about 5τ , then the above analysis for a constant amplitude sine wave approximately applies and the PSD output at any time will be proportional to the carrier amplitude during that integration time. The net effect over a time scale longer than the integration time is that the signal will be demodulated and passed as a low-pass filtered version of the envelope [90]. Hence the envelope will be faithfully passed if there are no significant frequency components above the low-pass cutoff. Therefore, the cutoff frequency of the low-pass filter of PSD (inversely proportional to τ) limits the bandwidth detected in the TRLIF signal.

Conversely, if all significant AC components are above the cutoff, then only the DC component will be passed and the output will be proportional to the time-average of the envelope waveform. This fact implies that a lock-in amplifier set to a long time constant measures time-averaged LIF signal. This is the theoretical foundation of the benchmark comparing the time-averaged TRLIF profiles with the average LIF profile from a lock-in amplifier set to a long time constant. If the two time-averaged profiles do not agree, then the time-resolved data acquisition or analysis likely failed somehow.

4.3.2.4 Upper and Lower Bounds of the Time Constant

Ideally we would like to satisfy the following double inequality in order to sufficiently preserve time-resolved information while using phase-sensitive detection:

$$T_{ref} \ll \tau \ll T, \quad (4.34)$$

where $T_{ref} = 2\pi/\omega_{ref}$ is the period of the reference sine wave at a frequency of ω_{ref} and τ is the integration time constant. T is a time scale of interest over which we would like to resolve change in the TRLIF signal envelope $F(t)$. The left hand inequality follows from the requirement of phase-sensitive detection to average over more than one reference period to recover the DC signal while rejecting the $2\omega_{mod}$ ripple. The right hand inequality is a statement that time-resolved information is destroyed when averaging over a few time constants.

The output signal will not be proportional to the TRLIF envelope if the requirement is insufficiently satisfied. If the left side is poorly satisfied, then the ripple will be excessive and obscure the DC signal proportional to the envelope amplitude. If the right side is poorly satisfied,

then features will be "smoothed over" because PSD is averaging data over a relatively long time. Or, in the frequency domain this means that the low-pass filter in PSD is removing significant frequency components of the envelope $F(t)$.

Consider an oscillation of about 10 kHz, such as the Hall thruster breathing and spoke modes. A measurement of this oscillation requires a resolved time scale of order $T = 10^{-5}$ s, since we require at least several points inside each period to reasonably resolve the oscillation. The fastest reference frequency possible would be ideal to reduce $2\omega_{mod}$ ripple and since the noise spectral density tends to decrease with frequency. Unfortunately, the AOM can reach at most less than 10 MHz, and the SNR is found to reach a maximum around 1 to 2 MHz (order $T_{mod} = 10^{-6}$ s). These bounds leave little room for the time constant, and apparently it will not easily satisfy Equation (4.34). We require a more careful analysis to determine appropriate bounds on τ .

The requirement that the low-pass filter rejects the $2\omega_{mod}$ ripple implies that the time constant is ideally larger than the modulation period. We can define a minimum allowable ratio of τ/T_{mod} that results in a maximum allowable ratio of ripple amplitude to DC amplitude. Since the filter gain for the DC signal is unity for any time constant, the ratio of output ripple amplitude to the DC amplitude is simply the amplitude gain at the $2\omega_{mod}$ frequency. We require the amplitude gain to be smaller than some maximum threshold gain $G_{2\omega}$:

$$G_{2\omega} > G(f = 2f_{mod}) = \frac{1}{\sqrt{1 + \left(\frac{f}{f_c}\right)^2}} = \frac{1}{\sqrt{1 + \left(\frac{4\pi\tau}{T_{mod}}\right)^2}}. \quad (4.35)$$

This leads directly to a minimum allowable ratio of τ/T_{mod} :

$$\frac{\tau}{T_{mod}} > \frac{1}{4\pi} \sqrt{\frac{1}{G_{2\omega}^2} - 1}. \quad (4.36)$$

The minimum allowable ratio $\min(\tau/T_{mod})$ is given for a few values of $G_{2\omega}$ in Table 4.1. A common standard for phase-sensitive detection in general is that the time constant is at least 10 reference periods, leading to a negligible ripple under 1%. In the case of TRLIF, the SNR is likely to be poor enough that the noise will obscure a ripple of even a few percent, making such a stringent requirement unnecessary. Apparently by this requirement, τ must only be a small factor larger than the modulation period; a time constant of about 1 to 2 times larger than T_{mod} should work.

Table 4.1. Select values of the minimum allowable ratio of time constant to modulation period $\min(\tau/T_{mod})$ given values of $G_{2\omega}$, a threshold of the maximum ripple amplitude to DC amplitude in the phase-sensitive detection signal.

$G_{2\omega}$	$\min(\tau/T_{mod})$
0.1	0.7918
0.05	1.5896
0.01	7.9573

The TRLIF signal envelope must be passed with little distortion, i.e. little difference in the gain of the relevant frequency components. Conversely to the previous calculation, a maximum allowable frequency component in the TRLIF envelope can be defined by imposing some minimum threshold amplitude gain G_{min} by the low-pass filter on the maximum relevant TRLIF signal frequency component f_{max} :

$$G(f = f_{max}) = \frac{1}{\sqrt{1 + \left(\frac{f_{max}}{f_c}\right)^2}} > G_{min}. \quad (4.37)$$

This requirement leads to an upper bound on the maximum relevant frequency component of:

$$f_{max} < f_c \sqrt{\left(\frac{1}{G_{min}^2} - 1\right)} = \frac{1}{2\pi\tau} \sqrt{\left(\frac{1}{G_{min}^2} - 1\right)}. \quad (4.38)$$

Putting the two requirements together results in a more precise inequality on τ :

$$\frac{1}{4\pi} T_{mod} \sqrt{\frac{1}{G_{2\omega}^2} - 1} < \tau < \frac{1}{2\pi f_{max}} \sqrt{\left(\frac{1}{G_{min}^2} - 1\right)}. \quad (4.39)$$

Note that the lower bound depends on the modulation period and a gain threshold, while the upper bound depends on the maximum relevant frequency component in the TRLIF envelope and the other gain threshold. The "maximum relevant frequency component" may not be well defined if the envelope is not sharply peaked in the frequency domain (which may be true for many conditions of a Hall thruster), in which case f_{max} can be defined empirically by testing a number of different time constants to detect where significant distortion occurs (see Subsection 5.4.4). The condition is still useful to conceptually determine the approximate range of appropriate time constant values.

These conditions to choose an appropriate value for the time constant are visualized in Figure 4.9 with $G_{2\omega} = 0.05$ and $G_{min} = 0.95$. The blue curves depict the gain of the $2\omega_{mod}$ ripple

for modulation frequencies of 1 MHz (dashed curve) and 2 MHz (dotted curve). The vertical lines indicate the minimum acceptable time constant from the condition of Equation (4.36). The red curve shows the condition on the maximum relevant frequency component from Equation (4.38). Frequencies below the curve satisfy the gain requirement G_{min} .

The intersection of the vertical lines with the red curve indicate the maximum allowable signal frequency component at the minimum acceptable time constant for each modulation frequency. This demonstrates how raising the modulation frequency allows the acquisition of higher signal frequencies. Thus we see that, given the threshold values used, the bandwidth of the system is about 35 kHz for $f_{mod} = 1$ MHz and 65 kHz for $f_{mod} = 2$ MHz.

The maximum acceptable time constant occurs where the red curve crosses below f_{max} , the maximum relevant frequency component of the TRLIF signal (not explicitly shown on the plot since it depends on the physical system under study). For example, the maximum acceptable time constant for a 10 kHz signal would be $\tau = 5.2 \mu s$.

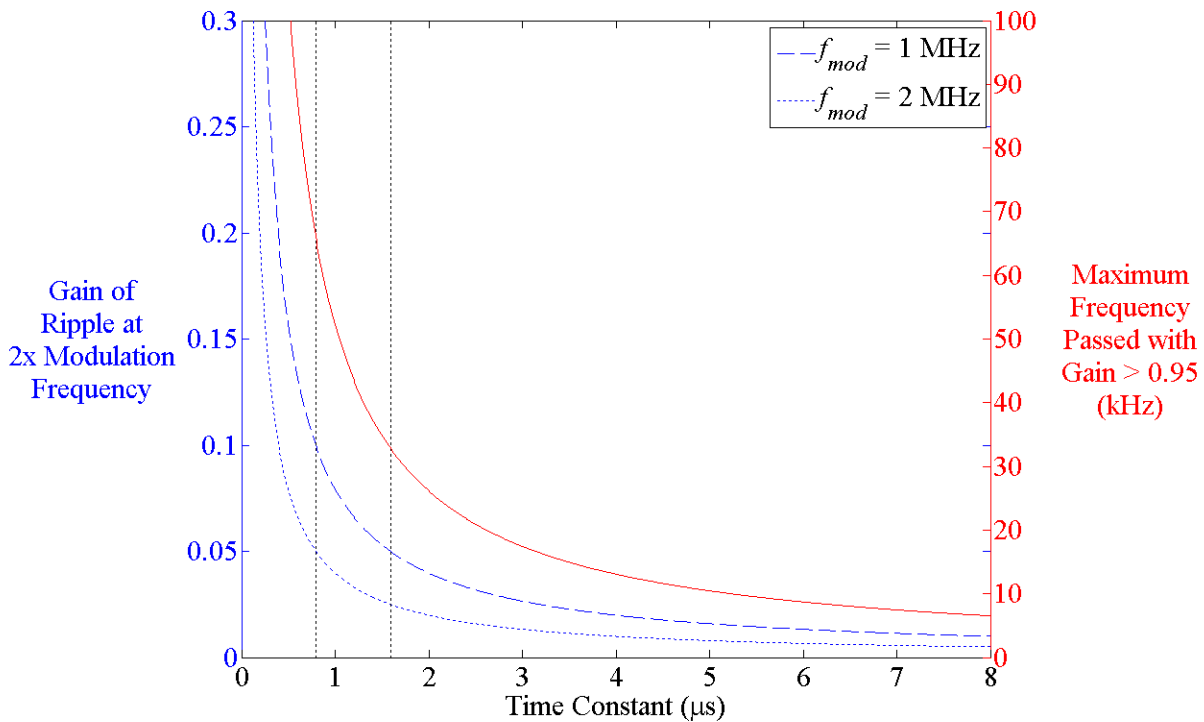


Figure 4.9. Gain of the $2\omega_{mod}$ ripple for 1-MHz and 2-MHz modulation frequencies (blue dashed and dotted lines) plotted together with the maximum frequency with gain exceeding $G_{min} = 0.95$. Both are functions of the time constant τ . The ripple gain determines the minimum acceptable τ while the maximum frequency passed determines the maximum acceptable τ .

Figure 4.9 shows the bounds of acceptable time constant values, but the improvement factor is proportional to the square root of the time constant. It would therefore generally be advisable to use the largest time constant possible to raise the SNR. Capturing high-frequency signals may require a low time constant, which may require a longer acquisition time to collect sufficient data to perform more averaging after the phase-sensitive detection stage. Thanks to the square root dependence, the difference is limited. A relatively large change of factor of four in the time constant only results in a change of a factor of two for the SNR improvement, which may not make a major difference for the final SNR while allowing a wide range of bandwidth.

The time constant is bound to the order of microseconds for the TFLIF technique to work. This time constant is about a factor of 10^5 to 10^6 smaller than typically used for time-averaged LIF experiments, and therefore it provides only a modest improvement to SNR because the SNR improvement factor is proportional to the square root of the time constant. Therefore, further averaging is necessary to recover the TRLIF signal from the noise after the PSD stage.

The final step in analysis is one of two averaging techniques. Averaging can be done over either the triggered time-domain PMT signal after filtering and phase-sensitive detection (called "triggered averaging"), or over transfer function estimators relating the PMT signal to some other input signal (called "transfer function averaging"). Both techniques, the topics of the following two subsections, are used with the same dataset to compare with each other for benchmarking. The final averaging step is the signal processing step responsible for the largest SNR improvement factor.

4.3.3 Fourier Analysis

4.3.3.1 The Assumption of a Linear System and the Existence of a Transfer Function

The main assumption in our use of Fourier analysis is that the thruster acts as a linear system in that there is a linear operator mapping some "input" signal (e.g. discharge current) to some "output" signal (e.g. LIF). If true, then knowledge of how the system maps input to output and a measurement of the input signal are sufficient to calculate the output. That is the strategy of TFLIF. The TRLIF signal is not measured directly; the transfer function and the input signals are measured, and the characteristic TFLIF output signal can then be calculated.

To avoid possible confusion with the many definitions of the word "linear" in other contexts, let us explicitly state here what it does **not** mean in this context. It does not mean that

the output is related to the input by a linear equation $y(t) = m \times x(t) + b$. Nor does it necessarily imply a high linear correlation coefficient. Nor is the cross correlation necessarily high at any time delay.

We assume only that there is a time-invariant linear system in the sense of the use of the term in linear algebra or systems and signals. Suppose O is the operator that describes the relationship between the input and output of the system. Given two time-varying input signals $x_1(t)$ and $x_2(t)$, suppose the system has output signals $y_1(t)$ and $y_2(t)$ such that

$$\begin{aligned} y_1(t) &= O\{x_1(t)\} \\ y_2(t) &= O\{x_2(t)\}. \end{aligned} \tag{4.40}$$

Then the system is linear if and only if it satisfies the following for any two input signals and constants α and β :

$$O\{\alpha x_1(t) + \beta x_2(t)\} = \alpha y_1(t) + \beta y_2(t). \tag{4.41}$$

In other words, if we take a linear combination of the original input signals to construct a third input signal, then the output of O associated with the new input signal is the same linear combination of the original output signals.

Consider discrete signal vectors $I_D[n]$ and $F[n]$, which are ideal discharge current and fluorescence signals in the absence of noise. This ideal fluorescence signal is proportional to the density of a population of ions in the interrogation zone having a velocity corresponding to the Doppler-shifted laser wavelength. Since this stage follows phase-sensitive detection, the ideal fluorescence signal $F[n]$ is a discrete, low-pass filtered version of the envelope $F(t)$ in Equation (3.3). Both $I_D[n]$ and $F[n]$ are discretely sampled a total of N times with n as the index of a particular sample, so $0 \leq n \leq N-1$. If the system is time-invariant and linear, then it can be shown that there is a constant transfer function relating the Fourier spectra of any two simultaneous input and output signals:

$$\tilde{F}[k] = H[k] \tilde{I}_D[k], \tag{4.42}$$

where k is the index for the Fourier space, which has the same range as n . It is our convention in this dissertation to use a tilde to denote the discrete Fourier transform, i.e. any signal $A[n]$ is related to $\tilde{A}[k]$ by the discrete Fourier transform (DFT):

$$\tilde{A}[k] = \sum_{n=0}^{N-1} A[n] \exp(-2\pi i k n / N). \quad (4.43)$$

A linear system relating input and output signals is a much more general relationship than the other meanings of the word "linear" mentioned above. It is possible for a linear system to map input to output such that output has different phase and relative frequency content from the input, though any frequency present in the output must be present in the input to some degree. It is also much looser than the assumption of periodicity used in the other TRLIF techniques discussed in Section 3.3. The assumption underpinning all of those techniques is that the physical process is periodically repeatable. In this analysis, the input and output waveforms may change and never be repeated while retaining the linear relationship.

In fact, the assumption is entirely based on the relationship between the input and output signals, not on the properties of the oscillations themselves. Since the characteristic output signal can be calculated for an arbitrary input signal, it is theoretically possible to capture even chaotic or stochastic oscillations that vary in some unpredictable way as long as the relationship between the input and output remains defined by the transfer function. Of course, the transfer function cannot capture oscillations where both input and output signals vary unpredictably and independently of one another, in which case it is clearly impossible to define a map between the two signals. It can, however, capture oscillations where the input signal varies chaotically or stochastically and the output signal is determined by the linear response of the system to the input signal.

As long as the relationship between the input and output remains time-invariant and linear then there is a transfer function relating the two signals. Then it is possible to measure the transfer function and calculate the characteristic output for any input signal. That characteristic output converges to the actual TRLIF signal as the average transfer function estimator approaches the true transfer function of the system.

4.3.3.2 Empirical Transfer Function Estimator

The simplest and most intuitive transfer function estimator is called the empirical transfer function estimator in the literature. Lobbia used this estimator in research on high-speed Langmuir probe measurements in Hall thrusters [36]. As discussed above, the idea is to assume the thruster acts essentially as a linear, time-invariant system in the sense that the TRLIF signal at a point in space is related to some input signal (e.g. discharge current) by a linear system.

Though some plasma behavior is stochastic and therefore necessarily nonlinear, linear characteristic features can exist and dominate behavior.

The assumption of a linear system relating Langmuir probe data and discharge current was ultimately justified by Lobbia by noting that the synthesized signals from the average transfer function matched all the key features of the original density signals for the same input current signal [37]. This comparison between the raw data and the characteristic signal from the average transfer function was possible because the raw data had a reasonable SNR, but it is not possible with TRLIF data since the raw signal is buried in noise. Another method of justifying the assumption of linearity that will be used for TFLIF is the benchmark proposed in Section 4.3.5.

Now, consider the measured signals, which we model to be composed of the ideal signals plus an additive noise sequence of random variables with some probability distribution with zero mean (since this signal is following filtering and phase-sensitive detection). The fluorescence signal is buried in noise, while the discharge current signal from the current probe is very precise:

$$\begin{aligned} F_{measured}[n] &= F[n] + N_f[n] \\ I_{D,measured}[n] &= I_D[n] + N_I[n] \approx I_D[n]. \end{aligned} \quad (4.44)$$

Then we can estimate the transfer function by simultaneously measuring the input and output signals and dividing their Fourier spectra elementwise:

$$\begin{aligned} H_{ETFE}[k] &= \frac{\tilde{F}_{measured}[k]}{\tilde{I}_{D,measured}[k]} \approx \frac{\tilde{F}[k] + \tilde{N}_f[k]}{\tilde{I}_D[k]} = H[k] + N_H[k] \\ N_H[k] &= \frac{\tilde{N}_f[k]}{\tilde{I}_D[k]}. \end{aligned} \quad (4.45)$$

$H_{ETFE}[k]$ is called the empirical transfer function estimator (ETFE). It is also buried in noise since the fluorescence signal is. During the experiment, we simultaneously digitize a series of N points of discharge current $I_{D,measured}[n]$ and the band-pass filtered photomultiplier signal, which becomes $F_{measured}[n]$ after phase-sensitive detection, for a total duration of about 60 s at a sampling frequency of about $f_s = 30$ MHz with the laser set to a constant wavelength. The 60-s series of N points is split into Q sub-series of N/Q points each. An empirical transfer function

estimator is found for each sub-series and they are all averaged together to obtain an average ETFE:

$$\langle H[k] \rangle = \frac{1}{Q} \sum_{q=1}^Q H_{ETFE,q}[k] = \frac{1}{Q} \sum_{q=1}^Q H[k] + N_{H,q}[k] = H[k] + \frac{1}{Q} \sum_{q=1}^Q N_{H,q}[k]. \quad (4.46)$$

Note that $N_{H,q}[k]$ is a complex random variable sequence with zero mean since $\tilde{N}_{f,q}[k]$ is and they are related by a linear transformation with no additive constant. $\tilde{N}_{f,q}[k]$ has zero mean since $N_{f,q}[n]$ does and they are related by the discrete Fourier transform, which is a linear transformation. Then by well-known properties of sums and linear transformations of random variables, the sum in the second term in Equation (4.46) is itself a sequence of complex random variables with zero mean and variance reduced by a factor of \sqrt{Q} relative to the average variance of $N_{H,q}[k]$. This is an example of the classic result that in general averaging Q measurements improves SNR by a factor of \sqrt{Q} . This simplistic analysis implies that **the average transfer function estimator converges to the exact transfer function** in the limit that Q approaches infinity.

The average transfer function is characteristic of the linear behavior of the thruster. Approximate values of $Q = 10$ or $Q = 100$ were used to average out turbulence and noise in Lobbia's work. Transfer function averaging is the primary form of averaging to improve SNR in a TFLIF measurement, hence values on the order of $Q = 100,000$ are used.

4.3.3.3 Advanced Transfer Function Estimators

The average ETFE in Equation (4.46) is the most intuitive estimator possible, following directly from the definition of a transfer function. There are other possible estimators and there is actually a significant collection of published research related to transfer function estimators that dates back to at least the mid-twentieth century. Casting the ETFE of Equation (4.45) into a more general notation common in the literature, we have:

$$\hat{g}_i = \frac{y_i}{u_i}, \quad (4.47)$$

where y_i and u_i are the DFT of the measured system output and input, respectively.

The problem with average of the ETFE is that the ETFE actually has theoretically infinite variance, and consequently the expected number of inaccurate frequency components remains

constant as the sample size (the variable Q in this dissertation) increases [91]. There is, therefore, a limit to the accuracy of the resulting synthesized output signal, regardless of the quantity of data included in the average. The simplistic analysis of Subsubsection 4.3.3.2 is perhaps intuitively helpful but flawed because it assumed a finite variance. The average ETFE does not converge properly to the exact transfer function because the variance is in fact infinite. Intuitively, the issue arises because when the measured system input (true system input plus noise) happens to be near zero for a particular frequency component, then the ETFE for that frequency component can be very large. The large and inaccurate instance of the ETFE then imposes a disproportionate influence that skews the average.

A significant body of research has focused on considering different transfer function estimators and averaging schemes. For example, the Welch estimator is a weighted average of the ETFE with the magnitude of measured input as the weighting factor:

$$\hat{g}_w = \frac{\sum_{i=1}^Q \hat{g}_i |u_i|^2}{\sum_{i=1}^Q |u_i|^2}, \quad (4.48)$$

where \hat{g}_i is the i^{th} instance of the ETFE and u_i is the corresponding measured system input. This is considered the "classical" scheme to ameliorate the problems of the ETFE. It helps by reducing the weight given to an erroneous ETFE. Heath showed not only that the Welch estimator has finite variance, but that it is the optimal weighting for both deterministic or stochastic inputs, two common models of input used to analyze transfer function estimators [92]. Deterministic input means the input is assumed to have definite values and the measured input varies only by random noise that is modeled as a zero mean complex Gaussian random variable. Stochastic input is a model for input where the input frequency components themselves are modeled as random variables with complex normal distribution.

Other exotic estimators and averaging schemes have been analyzed and can be useful in certain cases. For example, one scheme is to simply exclude from the average any instances of the ETFE inside a small region around the singularity, and another example is averaging the logarithm of the ETFE [93]. These have been shown to have value as well, but are more complicated to implement than the Welch average.

4.3.3.4 Synthesizing the Characteristic Fluorescence Signal

Suppose that the transfer function exists and has been measured by averaging over a large data set, yielding an average transfer function $\langle H[k] \rangle$. Then we can synthesize the characteristic response $F_{characteristic}[n]$ of the system to any input signal of our choice $I_D^*[n]$. Given an average transfer function $\langle H[k] \rangle$, the characteristic response of the system to the input is:

$$\tilde{F}_{characteristic}[k] = \langle H[k] \rangle \tilde{I}_D^*[k], \quad (4.49)$$

and we transform back to the time domain by an inverse DFT on the characteristic spectrum:

$$F_{characteristic}[n] = \frac{Q}{N} \sum_{k=0}^{N-1} \tilde{F}_{characteristic}[k] \exp(2\pi i k n Q / N). \quad (4.50)$$

Note that there is a separate transfer function for every spatial location and laser wavelength, and the output signals of each one correspond to the TRLIF signal of a velocity group at that point. The characteristic output of each transfer function is interpreted as the VDF amplitude as a function of time within the interrogation zone at a velocity corresponding to the Doppler-shifted ion absorption wavelength. The same synthesis procedure is repeated for all wavelengths desired to build up the time-resolved ion VDF profile at a single point in space. At least about 10 wavelengths are necessary to reasonably resolve the VDF in velocity space.

It is important to determine how the characteristic TRLIF signal can be interpreted. First, consider an arbitrary discharge current measurement $I_D^*[n]$ of length N/Q . If the assumption of linearity holds, then the average transfer function converges to the exact system transfer function as described in Subsubsection 4.3.3.2. Then by Equation (4.42) **the characteristic LIF signal spectrum converges to the exact spectrum of the LIF signal corresponding to the same time as the input $I_D^*[n]$** , and therefore $F_{characteristic}[n]$ (the inverse DFT) converges to the time domain fluorescence signal corresponding to that time. Therefore, the output of the linear model is the actual TRLIF signal at the time of the input discharge current trace, if the assumption of a linear system is justified.

It is important to note that the same input discharge current signal, $I_D^*[n]$, is used to synthesize the characteristic signal for all wavelengths. This is so that the synthesized response at each wavelength corresponds to a common input signal, ensuring coherent responses at all outputs regardless of how the input may randomly vary. From the discussion of the last two paragraphs, the characteristic output signals are regarded as the TRLIF signal at each laser wavelength that occurred at the time of the input trace. This is an important feature of the TFLIF

technique. The raw data series at each wavelength are all incoherent with each other since they are taken with some arbitrary time delay between them and the thruster oscillates chaotically. The TFLIF technique requires no triggering to measure the transfer function, yet signals synthesized from the same input are all coherent and can be meaningfully plotted together in a visualization of the VDF.

Note that the transfer function can be used to calculate a characteristic output signal for any input signal. The input $I_D^*[n]$ can be an arbitrary signal of length N/Q . The characteristic output will only be physically meaningful, however, if the input physically occurred. Furthermore, Lobbia pointed out that the characteristic output signal may only be accurate for input signals containing similar frequency components as the input/output signals originally used in the average to generate the average transfer function [37]. The average transfer function is likely to have large uncertainty at frequencies that were not well represented in the input/output signals used in the average. For a thruster operating in thermal equilibrium with constant operating conditions, chunks of input signal of length N/Q are all similar in terms of power spectral density. Since they are all similar, the particular trace used for the input does not make a major difference, so the last discharge current trace used in post-processing is often chosen.

4.3.4 Triggered Average

Following phase-sensitive detection, the signal is demodulated and following its original envelope, which corresponds to the population of ions in the interrogation volume with the velocity associated with the laser wavelength. We again model this measured signal as composed of the sum of the ideal TRLIF signal and some noise signal, which is assumed to be a sequence of independent random variables distributed by some probability distribution function with zero mean and some variance σ_F^2 . The noise will have zero mean at this stage because noise signals near the modulation frequency that are passed by PSD will have randomly varying phase and therefore have randomized sign.

As an alternative to transfer function averaging, we use triggered averaging as part of the validation argument. A triggered average takes an ensemble of traces, each measured after an instance of some trigger signal, and averages them all together elementwise (i.e. all of the same time points are averaged together). The main implicit assumption of triggered averaging is that there is a repeatable process that reoccurs at each trigger, thus we can consider the ideal LIF signal at the n^{th} time point after the q^{th} trigger to be composed of the average signal and some

fluctuation that may change due to small differences in the processes that occur after each trigger event. If the assumption that the same process reoccurs is exactly true, then the fluctuation is zero for all traces; but in general this is not true. The measured signal is modeled as:

$$F_{measured,q}[n] = F_{av}[n] + F_{fluc,q}[n] + N_{f,q}[n], \quad (4.51)$$

where $F_{av}[n]$ is the average TRLIF signal, $F_{fluc,q}[n]$ is some small fluctuation in LIF signal from the average, and $N_{f,q}[n]$ is the noise sequence. The index of the time bin is n , and the subscript q indicates that the signals occur on the q^{th} trigger. This is equivalent to the model given in Equation (4.44, with the only differences being the inclusion of the subscript q to denote different traces for each trigger and writing the ideal fluorescence signal as $F[n] = F_{av}[n] + F_{fluc}[n]$.

The idea of a triggered average scheme is to find triggers in the phase of the discharge current (or other reference signal) that indicate the phase of the repeatable process so that we can average the TRLIF signal traces at the same phase to recover the average signal waveform:

$$\langle F_q[n] \rangle = \frac{1}{Q} \sum_{q=1}^Q F_q[n] = \frac{1}{Q} \sum_{q=1}^Q F_{av}[n] + F_{fluc,q}[n] + N_{f,q}[n] \approx F_{av}[n] + \frac{1}{Q} \sum_{q=1}^Q N_{f,q}[n], \quad (4.52)$$

where $F_{fluc,q}[n]$ averages to zero in the limit $Q \rightarrow \infty$ since by definition the average LIF signal is $F_{av}[n]$. We split the modeled signal into time-averaged and fluctuation components to highlight that the measured TRLIF signal $\langle F_q[n] \rangle$ converges to the actual TRLIF signal after the triggers if and only if the assumption of a repeatable process is valid (i.e. that fluctuations from the average signal after each trigger are negligibly small).

$N_{f,q}[n]$ is a sequence of random variables, so the average of Q such sequences is another sequence of random variables with the same mean ($\mu=0$) and a variance reduced by a factor of Q :

$$\langle F_q[n] \rangle \approx F_{av}[n] + R(\mu=0, \sigma^2 = \sigma_f^2 / Q)[n]. \quad (4.53)$$

Ignoring the small contribution from the random fluctuations of the TRLIF signal from the average, the SNR improvement for this simple averaging process is the classic result for averaging Q measurements:

$$I_{TE} = \frac{SNR_{TE}}{SNR_{PSD}} = \frac{f_{av}[n]}{\sigma_f / \sqrt{Q}} \frac{\sigma_f}{f_{av}[n]} = \sqrt{Q}. \quad (4.54)$$

The triggered average has two applications for our purposes. First, in the case that we have both a linear system and a repeatable process that begins at some trigger, then both of the transfer function and triggered average techniques converge to the actual LIF signal, and therefore we can compare measurements of the two techniques to validate the TFLIF system. That is the idea behind the experiment with a periodic cathode oscillation in Section 6.3. Secondly, the triggered average is used for the benchmark in Subsection 4.3.5 since it can be shown that the transfer function can reproduce the triggered average even if the triggered average is not physically meaningful. This benchmark is demonstrated for a random cathode oscillation in Section 6.4

Note that the use of the triggered averaging technique following filtering and PSD is equivalent to the technique Scime et al. developed employing a lock-in amplifier and triggered averaging by an oscilloscope (see Subsubsection 3.3.1). The difference is that in this case the experiment is somewhat more flexible but has much more overhead because the analysis is done in software in post processing, whereas in that case PSD and averaging were done in real time with a commercial lock-in amplifier and oscilloscope.

4.3.5 General-Purpose Benchmark

4.3.5.1 Motivation for a New Benchmark

To be clear, the term "benchmark" is used here in the same way as it is commonly used in validating code for physical simulations or sometimes used in advanced diagnostics development. It refers to a special case where a known solution exists that can be compared with the results from the simulation or the new diagnostic to validate its accuracy. It is a consistency check that provides evidence that the new system works as intended, but does not provide a logically deductive argument that the system has no issues. There may be issues that for some reason did not surface in the case of the benchmark test. Successfully passing a number of benchmark tests may give researchers more confidence in the system, but of course the system can never be proven to be "perfect" by such tests. Therefore, it is important to keep in mind the possible fallibility of the system and to compare with other high speed diagnostics when possible. Using many different high speed diagnostics will, of course, give better insight into the physical processes at work in the thruster and also, critically, a better sense of whether some of the diagnostics may not be returning accurate results.

Although the validation tests based on periodic oscillations in Section 6.3 are positive, they do not necessarily lend much confidence to the technique when used with other plasma sources with nonperiodic oscillations. It would be desirable to have a general-purpose benchmark to test the results of the new technique under the circumstances of any arbitrary measurement.

One possible benchmark, in principle, would be to compare results on the same experiment with another TRLIF technique. Though possible, this presents logistical challenges since no single laboratory is equipped to perform two TRLIF techniques, nor do any two laboratories capable of different TRLIF techniques have copies of the same plasma sources. In addition, the main advantage of the PEPL TRLIF system is that it enables probing of nonperiodic oscillations that are not accessible with other TRLIF techniques anyway, thus such a test could not validate the system in the desired cases.

A general-purpose benchmark is proposed here that can be used to help validate the results for an arbitrary transfer function measurement, and therefore also validate the assumption of a linear map between the input and output of an arbitrary system of interest. It can be shown that if the system of interest is linear, then the transfer function is theoretically capable of reproducing the result of the triggered average. Therefore, the results of those two independent analysis techniques used on the same data set can be compared as a benchmark that helps to validate the transfer function.

4.3.5.2 The Reproduction of the Triggered Average by the Transfer Function

Most of the time, we will be interested in calculating the characteristic output for an input trace that physically occurred, since Subsection 4.3.3.4 argued that the characteristic output then corresponds to the actual output signal that occurred at the time of the input signal. But it is important to remember that the transfer function can be used to calculate a characteristic output signal for any possible input signal of length N/Q . If the input signal did not physically occur then, of course, the characteristic output signal will not correspond to an actual output that physically occurred.

Nonetheless, consider the characteristic TRLIF signal calculated from the transfer function using the triggered average of discharge current traces as the input to the transfer function:

$$I_D^*[n] = \frac{1}{Q} \sum_{q=1}^Q I_{D,q}[n]. \quad (4.55)$$

If the system is time-invariant and linear, then the transfer function mapping input to output exists and Equation (4.42), the formula for calculating output signal, is valid. Insert the input signal above into the formula for characteristic output signal and convert to the time domain by applying the inverse discrete Fourier transform to both sides:

$$\begin{aligned} \tilde{F}_{characteristic}^*[k] &= \tilde{I}_D^*[k] \langle H[k] \rangle \\ F_{characteristic}^*[n] &= \mathcal{F}^{-1} \left(\tilde{I}_D^*[k] \langle H[k] \rangle \right) \\ F_{characteristic}^*[n] &= \mathcal{F}^{-1} \left(\mathcal{F} \left(\frac{1}{Q} \sum_{q=1}^Q I_{D,q}[n] \right) \langle H[k] \rangle \right). \end{aligned} \quad (4.56)$$

Note that the discrete Fourier transform can move inside the sum and act on each individual term in the sum because the discrete Fourier transform is a linear transform:

$$\begin{aligned} F_{characteristic}^*[n] &= \mathcal{F}^{-1} \left(\frac{1}{Q} \sum_{q=1}^Q \mathcal{F} (I_{D,q}[n]) \langle H[k] \rangle \right) \\ F_{characteristic}^*[n] &= \mathcal{F}^{-1} \left(\frac{1}{Q} \sum_{q=1}^Q \tilde{I}_{D,q}[k] \langle H[k] \rangle \right). \end{aligned} \quad (4.57)$$

Now, recognize that each term in the sum in fact corresponds to the right hand side of the formula for characteristic output signal of each of the q traces in the triggered average:

$$F_{characteristic}^*[n] = \mathcal{F}^{-1} \left(\frac{1}{Q} \sum_{q=1}^Q \tilde{F}_{characteristic,q}^*[k] \right). \quad (4.58)$$

The inverse discrete Fourier transform is also a linear transform, thus it can move inside the sum and act to convert terms to the time domain:

$$F_{characteristic}^*[n] = \frac{1}{Q} \sum_{q=1}^Q \mathcal{F}^{-1} \left(\tilde{F}_{characteristic,q}^*[k] \right) = \frac{1}{Q} \sum_{q=1}^Q F_{characteristic,q}^*[n]. \quad (4.59)$$

It was argued in Subsection 4.3.3.4 that, if the system is linear, the characteristic output signal converges to the actual TRLIF signal that occurred during the time of the input trace. That means in this case that the $F_{characteristic,q}^*[n]$ corresponds to the TRLIF signal that occurred during

the q^{th} discharge current trace. If so, then average on the right hand side converges to the triggered average of TRLIF signal if the system is indeed linear:

$$F_{characteristic}^*[n] = \frac{1}{Q} \sum_{q=1}^Q F_{characteristic,q}[n] \xrightarrow{(Q,N) \rightarrow (\infty,\infty)} F_{av}[n], \quad (4.60)$$

where $F_{av}[n]$ is the triggered average of TRLIF signal from Subsection 4.3.4. We have shown that the characteristic output of the average transfer function will theoretically reproduce the triggered average TRLIF signal if the triggered average discharge current is used as the input signal and the system is indeed linear.

4.3.5.3 Significance of the Benchmark

The result of Subsubsection 4.3.5.2 suggests a possible benchmark. The "reproduction" of the triggered average TRLIF signal that is calculated using the transfer function can be compared to the actual triggered average TRLIF signal to verify that calculations based on the average transfer function are accurate.

A direct comparison of the triggered average TRLIF signal with the characteristic TRLIF signal of an actual input discharge current trace is only valid in the special case of periodic oscillations, where both signals converge to the actual TRLIF signal. This special case for periodic oscillations is the concept behind the experiment in Section 6.3. The new concept for a benchmark is a significant development because it does not rely on the special case of periodic oscillations to make triggered averaging valid. Instead, the transfer function can reproduce the triggered average signal whether the triggered average is physically meaningful or not. This allows a direct comparison between the triggered average TRLIF signal and the characteristic signal of the transfer function using the triggered average discharge current trace as input. Thus the benefit of the proposed benchmark is that it is "general-purpose" in that it can be used for any arbitrary oscillation.

If the two signals agree, as is theoretically expected if the system is linear, then this benchmark provides evidence that the Fourier analysis works as expected for the particular case tested. Specifically, the comparison simultaneously provides evidence that system of interest is indeed linear and that the parameters chosen in analysis (such as chunk size, windowing, and total length of the data set) are reasonable for the system of interest. The logic for this is that if those conclusions were incorrect then we would expect some significant error in the transfer

function's reproduction of the triggered average signal. Hence, if there is not significant error in the comparison then we can reasonably conclude that the system of interest is likely linear and that the analysis parameters are reasonable for the system of interest.

Chapter 5

A Practical Implementation of the TFLIF Technique

Experience without theory is blind, but theory without experience is mere intellectual play.

– Immanuel Kant

Building on the foundation laid out in the previous chapter, this chapter covers many of the details of the implementation of a practical TFLIF system at PEPL. The preceding chapter covered the theory of the signal processing technique. If the theoretical considerations were changed significantly, then the result could likely be considered a distinct measurement technique. This chapter focuses on the finer details of the hardware, experimental setup, and algorithms that define exactly how that system is implemented. These details could be changed and the underlying technique largely would remain the same, though different implementations could have vast differences in efficiency and reliability. In fact, this chapter represents a major thread of the research presented in this dissertation: to approach the best implementation possible, starting from the simplest and naivest implementation possible. Only after developing a system that reliably and efficiently makes measurements can we begin to validate the accuracy of those measurements in Chapter 6 and demonstrate useful measurements in a Hall thruster in Chapter 7.

The essential goal of this chapter is to recount the salient details of this progression so that the reader understands why certain choices were made and the general idea of how to implement a TFLIF system. This chapter is concerned with three broad areas. Firstly, hardware considerations are covered in Section 5.1. Secondly, there a number of important software considerations. The TFLIF technique would not be possible without a reliable phase-locking scheme (Section 5.2) and not practical without the savings from data recycling (Section 5.3). Other important software topics are included in Section 5.4. Finally, Section 5.5 presents the efforts to optimize the SNR and raise the operating modulation frequency to enable better signal bandwidths.

5.1 Hardware Considerations

TFLIF relies on hardware that was not readily available just a few years ago. It requires a high-speed digitizer capable of continuously streaming about 10 GiB of data to the PC, a storage system with a sufficiently fast transfer speed and large capacity to support the massive datasets, and the computational power to process the datasets in a practical time frame. TFLIF only became a practical possibility within the past several years as these features became available.

The goal of this section is not to present an exhaustive trade study that was done to select the particular hardware for the implementation of TFLIF at PEPL; that is already obsolete. Rather, the goal is to convey the important considerations in selecting and using appropriate hardware for a TFLIF system. The focus is especially on the requirements, limits, and optimal values of relevant parameters for TFLIF. Subsection 5.1.1 covers the digitizer and the main concerns of data transfer, sampling frequency, and voltage resolution. Subsection 5.1.2 discusses how the RAID is the only feasible solution for the data transfer and storage requirements. Subsection 5.1.3 considers the acousto-optic modulator (AOM) necessary for high-speed modulation. The other significant piece of hardware for TFLIF, the band-pass filter, has been covered in Subsections 4.3.1 and 5.5.7.

5.1.1 Digitizer

A dataset for a single laser wavelength continuously sampled in time is about 10 GiB (cf. Subsection 5.1.2), precluding acquisition to memory onboard the digitizer card since onboard memory is usually only in range of 64 MiB to a few gigabytes. Even if possible, acquiring to memory on the card and then transferring to permanent storage would add an extra step and waste time during the experiment. It is much more practical for this application to stream directly to permanent storage on the PC. At the time of building this system, Alazartech was the only digitizer manufacturer with PCI express models capable of streaming directly to PC memory or storage at the full sample speed. Though the feature is more common now, this was the deciding factor in selecting the digitizer at the time.

The following subsections discuss the requirements, limits, and optimal settings on the parameters of the digitizer that are relevant to the TFLIF system. The requirements and limits on the sample frequency are discussed in Subsubsection 5.1.1.1, as well as an argument that efficiency could be improved with a more complete test of the SNR versus sample frequency.

The input voltage scale is the subject of Subsubsection 5.1.1.2, where the importance of matching the output band-pass filter signal to the input scale is emphasized.

5.1.1.1 Sampling Frequency

Since the modulation frequency must be on the order of megahertz (see Subsection 4.3.2), the sample speed f_{sa} must be at a minimum on the order of tens of megahertz to resolve the modulation waveform for phase-sensitive detection. This requirement is not difficult; the Alazartech ATS9462 digitizer selected for TFLIF is capable of a sample frequency up to 180 MHz. It is important to note, however, that the maximum sample speed is limited by the function generator providing the clock signal when operating with an external digitizer clock, and operating in the external clock mode is critical for this application (see Section 5.2). The Agilent 33522A, the function generator used to generate the external clock, is limited to 30 MHz.

To save on storage and computational expense, the optimal sample speed setting may be considered to be the lowest sample speed that still reliably captures the signal with no significant reduction in SNR. A test of the effect of the sample frequency in Figure 5.13 in Subsection 5.5.5 shows that the SNR is not consistently affected by the changing the sample frequency from 20 MHz to 30 MHz for modulation frequencies in the range of 0.1 to 3 MHz. Previous analyses using an older analysis code had shown a SNR slightly but consistently higher at 30 MHz than at 20 MHz for all modulation frequencies. The reason for the change is currently unknown, but could be related to a problem with the earlier code such as a less reliable phase-finding routine. Much of the data presented in this dissertation were captured at 20 or 30 MHz, but if the latest analysis is correct, then the sample speed may be higher than optimal. More work is necessary to determine the optimal ratio of sample frequency to modulation frequency.

5.1.1.2 Voltage Resolution

The voltage resolution of the digitizer is an important feature to consider. The ATS9462 has 16-bit resolution, meaning that voltage measurements are recorded as 16-bit integers and there are $2^{16} = 65536$ voltage levels uniformly spread between the positive and negative full scale voltage. The actual analog voltage value is rounded to the nearest voltage level resolved. Digitizers with higher sample frequency tend to have lower voltage resolution, and 16-bit resolution is generally the maximum available. 2^{16} levels may seem sufficient to well resolve the signal, but it is more difficult for a noisy signal. For example, the background light collected can

be about 1000 times the amplitude of the LIF signal. The SNR improvement factor of the band-pass filter is estimated to be about 50 [89], so the SNR of the digitized signal may be about 53:1000. In order to not destroy the signal by saturating the digitizer above/below full scale, the full scale must be set to at least about the scale of the background noise, leaving only about $65536 * 53/1000 = 3473$ levels near the full scale of the signal. The effective voltage resolution of the signal is therefore between 11-bit and 12-bit levels at best. Clearly, it is important to have the best voltage resolution possible in an application where the goal is to resolve a weak signal in a strong noise background.

The signal is not resolved extremely coarsely at this resolution, but reducing the effective signal resolution even further by a significant factor could have deleterious effects. The voltage quantization introduced by the digitizer is a source of noise that becomes increasingly significant as the effective signal voltage resolution is reduced, though it only strongly affects SNR at very low effective voltage resolution of a few bits [94]. The danger is that a careless choice of digitizer input scale to band-pass filter output voltage can annihilate the signal either by saturating the digitizer or by setting the voltage resolution much larger than the background signal scale.

To help ensure that the effective signal voltage resolution is not unnecessarily reduced, the best practice is to match the output voltage of the band-pass filter to the input scale of the digitizer. There are several parameters to change to make a good match. The digitizer full scale and input/output gains of the bandpass filter may be changed in a discrete set of selections to roughly match the output voltage to the input scale. The PMT bias voltage may then be adjusted to continuously change the voltage level since the current gain is a monotonically increasing function of bias. Quantum efficiency also increases as a function of bias. Therefore, the parameters should optimally be adjusted so that the PMT bias is as large as possible without exceeding the maximum (1000 V for the Hamamatsu R928 currently in the TFLIF system) to maximize the photoelectron rate.

No comprehensive test of how the matching affects SNR is available, but three points were tested early on when the system still used the internal digitizer clock. The signal at five different points was measured, but only three null signals were measured (for unknown reasons), enabling the SNR measurement at those points. The ratio of saturated to unsaturated points (R_{sat}) is measured by sampling the filtered PMT signal for about 1 s, adding together the number of

points saturated at the maximum and minimum full scale, and dividing by the total number of points sampled. The ratio was adjusted during the test by changing the photomultiplier bias between 864 V and 949 V while holding all other parameters constant. The results are summarized in Figure 5.1. Surprisingly, SNR increases somewhat as a function of R_{sat} in the range of values tested.

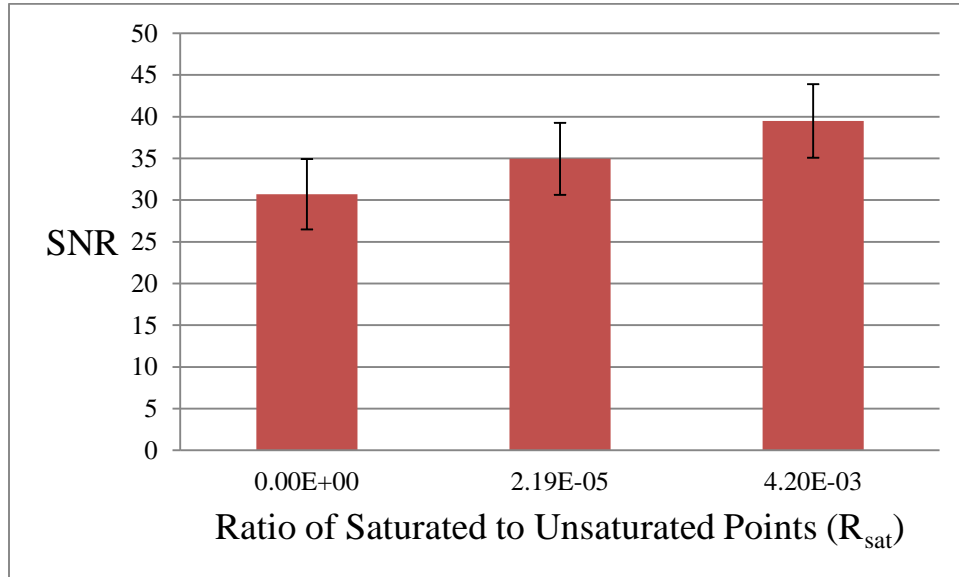


Figure 5.1. The SNR moderately increases as a function of R_{sat} .

This result suggests that the effective voltage resolution may actually be worse than estimate above because apparently the contribution of quantization noise may be significant. The SNR will surely have a maximum value before $R_{\text{sat}} = 1$ since at that point the digitized signal will be saturated at all samples and the signal lost. It may be possible to improve performance by operating at a higher value of the ratio, but it is unclear when the loss of signal from saturation will dominate the gains from a better effective signal voltage resolution. A more comprehensive test with a larger range of values is necessary to determine the optimal conditions.

Testing for distortion in the TFLIF signal due to the digitizer saturation in addition to measuring the time-averaged SNR would be a sensible precaution. Caution must be taken with such a test with a higher range of R_{sat} values to ensure that the maximum voltage encountered is safe for the digitizer electronics. In addition, operating at a high R_{sat} value in the long term would also increase the risk that some unexpected change would expose the digitizer to unsafe voltages, which may not be worth a modest SNR improvement unless some overvoltage protection is added to the system.

In the absence of a more comprehensive test and careful overvoltage protection, a general rule is that the match is sufficient when the fraction of saturated to unsaturated points is between about 10^{-6} to 10^{-4} . It is the middle of the range of R values tested and corresponds to at most a few saturated points in a chunk of size 0.001 s used to compute an empirical transfer function, which likely have negligible effect. The benefits of the rule are that the two-order-of-magnitude range is easily achievable, and it prevents either pitfall of the saturation or unnecessary reduction in the effective signal voltage resolution that would result from an extremely poor match.

5.1.2 Redundant Array of Independent Drives (RAID)

One of the most challenging requirements of the technique is the sheer volume of data that must be acquired. The dataset is collected at a rate of Bit Resolution \times Number of Channels \times f_{sa} bits per second. This is about 114.4 MiB/s for an acquisition with $f_{sa} = 30$ MHz and 2 channels (the configuration of the hollow cathode tests in Chapter 6) or about 171.7 MiB/s for the 3-channel configuration of the H6 campaign in Chapter 7. This leads to about 6.7 or 10.1 GiB respectively for a 60-s acquisition per wavelength. With a 60-s dataset for each of 20 laser wavelengths, for example, that totals to 134.1 or 201.2 GiB per VDF. Clearly a campaign containing a map of VDF measurements at many locations or operating conditions can easily reach into the terabyte range.

A Redundant Array of Independent Disks (RAID) was at the time of the equipment purchase, and as of writing still is, the only solution that satisfies these requirements for both capacity and transfer speed. Hard disk drives were available up to about 3 TB but are limited to between 100-200 MiB/s sustained write speed. The fastest solid state drive offered write speed up to 720 MB/s for the OCZ RevoDrive X2, but capacities were under 1 TB. The capacity and write speed of RAID, on the other hand, scales with the number of drives in the array as long as a high quality RAID controller efficiently manages the transfer.

Technical details of RAID are beyond the scope of this dissertation, but for completeness we briefly outline the salient ideas to appreciate how the RAID fits into the TFLIF system. A RAID takes data that would normally be consecutively written to a single drive (for example, a 10 GB data file) and splits it into many "blocks" of some standard size (often on the order of KiB or MiB) that are then written to the n drives in the RAID simultaneously. The unit of n blocks written to the n drives is called a stripe, and this process is called striping. A RAID can have a large data rate advantage over a single drive since the data are written to or read from many

drives simultaneously. A RAID is redundant since it can be set to write copies of the same information on different drives (mirroring) or to write "parity" information that can be used to recover the data in the event that some of the drives in the RAID fail. The RAID controller manages all of the overhead involved with calculating parity data and reading from/writing to the drives, and it presents the RAID to the OS as a single drive.

There are many different schemes defining how blocks are written to the drives, called RAID levels. The level RAID 6 was chosen as a good balance between robustness and storage efficiency. Two blocks of parity information are added to each stripe so that it can operate with no loss of data with up to two failed drives. Since two of the n blocks written to the drives simultaneously contain "extra" parity information, the write speed of the RAID is theoretically a factor of $n - 2$ faster than a single drive, and the effective capacity is a factor of $1 - 2/n$ times the sum of the capacity of the n drives.

The 10-drive RAID at PEPL, designed to exceed the transfer and storage requirements, is composed mostly of Western Digital Red drives² with a capacity of 2 TB and internal transfer rate of 147 MB/s. This leads to a net capacity of 16 TB (14.9 TiB) and a theoretical transfer rate of 1176 MB/s (1122 MiB/s). The RAID is capable of sustained streaming at the maximum sample speed for minutes at a time, whereas the digitizer has buffer overflow issues within seconds even with the OCZ RevoDrive X2, despite the nominal transfer speed matching the streaming requirement of 720 MB/s (686.6 MiB/s). The net capacity has proven enough for the several initial tests, validation campaigns with a hollow cathode, and the initial Hall thruster campaign. As of this writing, the RAID is about 75% full. More capacity will be necessary for more extensive TFLIF campaigns in the future, but fortunately hard drive technology has advanced significantly since building the RAID. RAID controllers supporting drives larger than 2 TB and 6-TB Western Digital Red drives are now readily available.

5.1.3 Acousto-Optic Modulator

While a mechanical chopper wheel is sufficient for a normal LIF experiment, there are drawbacks that make it unacceptable for a time-resolved experiment up to the bandwidth desired. The most significant is simply that choppers are limited to modulation frequencies on the order

² The RAID was initially built with Hitachi Deskstar drives with a higher transfer rate of 160 MB/s. Most of these have failed over the years and were eventually replaced with the Western Digital Red drives, which have proven to fail at a far lower rate. Unbeknownst to the author when originally building the RAID, the Hitachi Deskstar series has the popular but dubious appellation of "Deathstar".

of kilohertz, but Section 4.3.2 shows that the use of phase-sensitive detection requires modulation frequencies on the order of megahertz. Another potential issue even if they offered sufficient speed is that they do not output a stable frequency. The display shows random variation in the modulation frequency about a few percent of the setting. A lock-in amplifier can handle the instability if the reference is directly measured, but such variation could be difficult for the phase-locking necessary for TFLIF. In order to allow higher modulation frequency, we must use a more complicated instrument and sacrifice the ideal operating characteristics of the chopper. The maximum transmission is 100% because the laser simply passes through holes in the chopper wheel. The contrast extinction ratio, the ratio of maximum to minimum transmission, is infinite because the laser is completely blocked by the solid sectors of the wheel.

An acousto-optic modulator (AOM) provides a similar effect modulating the laser power in a much larger range of frequency. It uses a piezo actuator to drive an acoustic wave in a transparent crystal, creating periodic regions of higher and lower density that act as a grating for Bragg diffraction. Bragg diffraction normally occurs for x-rays incident on a crystal lattice, but the larger length scale of the acoustic wave is appropriate to Bragg diffraction of light in the IR range.

The maximum modulation frequency that the AOM can support is determined by the 10%-to-90% rise time, which is given by the formula:

$$t_r = \frac{1.3d_0}{2V}, \quad (5.1)$$

where d_0 is the diameter of the laser beam within the crystal and V is the acoustic velocity of the crystal. For the TeO₂ material of the NEOS model 23080-1 at PEPL, $V = 4260$ m/s. The rise time for the 1-mm laser beam is insufficient, so a lens focuses the beam to a smaller waist within the AOM and a second lens collimates the modulated beam following the AOM. For a Gaussian beam, the waist upon focusing by a lens is diffraction limited by:

$$d_0 = \frac{4f\lambda}{\pi d_1}, \quad (5.2)$$

where f is the focal length of the lens, λ is the laser wavelength, and d_1 is the beam diameter prior to focusing. The maximum frequency supported can be estimated by a period of four times the rise time, which allows for a rudimentary "square wave" with rise, fall, maximum, and minimum phases each taking one rise time.

Another consideration is intensity of the beam within the crystal. A high power laser focused to a small spot can damage the TeO₂ crystal, which has a maximum permissible intensity of 1000 W/mm. The average intensity of the beam is estimated by dividing a nominal 100 mW power by the area of a circle with diameter d_0 , where $d_1 = 1$ mm is an estimate of the beam waist before focusing into the AOM. The actual laser power can be more than 100 mW, but not more than about a factor of 4 greater. The maximum intensity for a Gaussian beam is a factor of two higher than the average.

Table 5.1 provides a summary of the operating characteristics of two sets of focusing/collimating lenses used at PEPL. Both sets of lenses produce a maximum intensity far below the maximum permissible intensity. In general, the intensity will not become a problem unless working with a higher power laser, another AOM designed for a smaller focus spot, or another AOM designed for coupling with a single-mode fiber. The rise time for this model of AOM is limited to that achieved with the 120-mm lenses due to the acceptance angle of the AOM. The diffraction efficiency as a function of beam diameter and rise time is given by Figure 1 in the NEOS 23080-1 manual, and the values were verified in the installation at PEPL.

Table 5.1. A summary of the operating characteristics of two sets of focusing/collimating lenses. There is a trade-off between the maximum frequency f_{max} and the diffraction efficiency η_{diff} .

	d_0 (mm)	t_r (ns)	f_{max} (MHz)	I_{max} (W/mm ²)	η_{diff}
$f = 200$ mm	0.21	32	7.7	5.6	~70%
$f = 120$ mm	0.12	19	12.8	15.6	~60%

The first set had a focal length of 200 mm, with a high diffraction efficiency of $\eta_{diff} = 70\%$ but marginal estimated maximum frequency of $f_{max} = 7.7$ MHz. This setup was chosen to maximize SNR while working out the bugs in the system at somewhat lower modulation frequencies. The borderline estimated maximum frequency implied that the setup may need to be changed to reach frequencies as high as several megahertz.

Eventually, evidence was detected by the photodiode that the laser power modulation waveform from the AOM becomes increasingly distorted at frequencies of only a couple of megahertz, despite the nominal rise time implying that frequencies above five megahertz may be possible. It was suspected that the distortion may contribute to the maximum in SNR versus modulation frequency that was shown in Figure 5.12.

The effect of the lenses on waveform distortion and the SNR in LIF measurement was tested by comparing with the results of a 120-mm set of lenses. The distortion was reduced, and the maximum does indeed occur at a significantly higher frequency with the 120-mm lenses, as shown by comparing Figure 5.12 with Figure 5.13. Though the maximum SNR occurs at a higher frequency, the SNR values are generally somewhat lower with the 120-mm lenses, as expected with a reduced diffraction efficiency of 60% instead of 70%.

Unfortunately, the rise time of the model 23080-1 is limited to at best about 20 ns due to the acceptance angle and electrical bandwidth. If the laser modulation waveform is still a limiting factor for the SNR vs modulation frequency, then another AOM model may be necessary.

The model 15210 was recommended as an option because it is designed for a shorter rise time less than 10 ns. The model 23080-1 was conservatively chosen to maximize SNR because it was not known for certain what the SNR of time-resolved data would be following analysis and how much data acquisition and averaging would be necessary. The 23080-1 features both higher maximum diffraction efficiency (85% vs. 70%) and static transmission (97% vs 95%). As the TFLIF system matures and SNR becomes less of a concern, a modest loss in SNR may be a reasonable trade-off for improved modulation frequency.

5.2 Phase Locking

5.2.1 Defining the Problem of Phase Locking for the TFLIF System

As explained in Subsection 4.3.2, the origin of the name of phase-sensitive detection is that the amplitude of the output signal is proportional to the cosine of the difference between the phase offsets of the input and reference signals:

$$V_{out} \propto \cos(\phi_{signal} - \phi_{ref}). \quad (5.3)$$

This fact complicates implementing phase-sensitive detection because the not only must the frequency of the signal be known, but the phase offset must also be known and matched by the reference signal used to recover the signal from the noise. If the reference signal is not held at the same phase as the signal, then the signal amplitude will be incorrect. It will either introduce erroneous time-resolved features if the reference drifts into and out of phase at a rate resolvable by the system, or will lead to a complete breakdown of the analysis with only a noise output

signal if the drift occurs at a fast rate (such as if the reference frequency does not quite match the signal frequency).

Phase-locking is not usually a problem for a typical lock-in amplifier because the reference signal is measured directly from the source of the signal modulation (e.g. an optical chopper) and it is used in real time with analog or specialized digital circuitry to recover the signal. Even if the physical system of the experiment changes the phase offset of the signal from the reference phase offset, dual-phase lock-in amplifiers use two phase-sensitive detectors with a 90° phase shift, where the output of each detector is commonly called X and Y. Since X and Y are 90° apart in the Argand plane, the X and Y signal amplitudes are added in quadrature to find the magnitude of the input signal regardless of phase:

$$Z = \sqrt{X^2 + Y^2} . \quad (5.4)$$

This scheme in lock-in amplifiers is known as a vector computer, and is very common to most commercially available, modern lock-in amplifiers, including the SR-810 lock-in amplifiers at PEPL.

Phase locking is a unique challenge to the TFLIF system because the reference is not measured directly but must be inferred from the signal itself. Instead of measuring the reference and using specialized instrumentation to recover the signal in real time, only the linearly related input and noisy output signals are sampled and saved for later analysis. The TFLIF system is akin to a lock-in amplifier without a direct reference measurement. It is, of course, technically possible to measure the reference signal, in this case it would be the AOM driving signal, but we decided to avoid it because it would require an extra digitizer channel, raising the already steep storage requirements considerably. Instead, the reference signal is simply a numerically generated waveform (usually sine) at the same sample frequency f_{sa} and chunk size as the dataset so that they can be multiplied together. In general, the reference signal $V_{ref}[n]$ is defined as:

$$V_{ref}[n] = \sin(\omega_{ref}[n]t[n] + \phi_{ref}[n]) , \quad (5.5)$$

where (with malice aforethought) we may allow the reference frequency and phase offset to change depending on the sample n within the dataset. The goal, then, is to find the frequency and phase offset to use to generate the reference signal. Note that the global dataset time is defined with $t[n = 1] = 0$ at the first sample in the dataset, and the phase offset in the reference is defined based on that time.

5.2.2 Difficulties in Finding the Optimal Reference Frequency and Phase Offset

The simplest and naivest method to handle the unknown phase is would be to mimic the action of a lock-in amplifier vector computer: simply have two reference signals 90° out of phase from each other, and define the demodulated output signal to be the sum of their outputs added in quadrature as in Equation (5.4). The problem with this scheme is that the signals are still very noisy following phase-sensitive detection with a short time constant ($\text{SNR} \ll 1$), and combining two such signals together by Equation (5.4) would reduce the SNR by propagation of error. It would require an impractical amount of averaging in the next step (transfer function or triggered averaging).

Instead, a workable solution is to approach the problem in two steps. First, determine the phase offset separately throughout the whole 60-s dataset. Second, apply phase-sensitive detection with the previously defined optimal reference signal while proceeding with transfer function averaging.

Unfortunately, there are a few details that complicate the task of finding the optimal reference signal. In principle, we know that the signal frequency equals the modulation frequency, which we set using the function generator driving the AOM, and therefore we only need to find the phase offset of the signal in the dataset. In practice, neither the modulation frequency f_{mod} (from the function generator driving the AOM) nor the sampling frequency f_{sa} (from the digitizer clock) are known to perfect precision or are even constant. The fairly standard Agilent model 33220A has a nominal frequency stability of 10-20 ppm. The Tektronix AFG3101, the best function generator in the laboratory when this project started, has a nominal frequency reference stability of 1 ppm. A higher precision model was later purchased in part due to the frequency stability issues identified in the TFLIF project. The Agilent 33522A has a nominal 0.1 ppm frequency stability with the high stability frequency reference option. Meanwhile, the stability of the internal sampling clock of the Alazartech 9462 is 25 ppm for sampling frequencies at or below 125 MS/s, which is the limiting factor in this case. The manufacturers recommend warming up the instruments for 30 minutes to reach the nominal stability, and the stability tends to worsen within about a year after the calibration. Not only are f_{mod} and f_{sa} unknown, but it is important that they are not necessarily constant either. The measured signal frequency tends to drift over a time scale on the order of seconds or minutes by

a few ppm, even after the 30-minute warm-up time. This drift is on a significant time scale for the time-resolved LIF measurement with datasets about 60 s long and needs to be corrected.

Both uncertainties in the sampling speed and in the modulation speed will affect the actual signal frequency observed in the sampled data. A stability as low as 1 or even 0.1 ppm may not sound like much, but even they can be a problem for phase-sensitive detection, let alone the 25 ppm stability of the digitizer clock. For example, a 1 MHz modulation signal measured with 0.1 ppm accuracy may actually be 1,000,000.1 Hz in the sampled dataset while the reference signal is defined to be 1 MHz. These are close enough that phase-sensitive detection may still detect the signal if it is in the transmission window defined by the output low-pass filter. Even so, the two signals would drift out of phase at a rate of $2\pi(f_{signal} - f_{ref})$, or a full period every 10 s. On the time scale of 5τ over which PSD operates, the signal and reference will appear to be separated by a constant phase difference, hence Equation (5.3) applies. With a 60-s long dataset, the reference signal would drift out of and back into phase with the signal six times over the course of the dataset, and the amplitude of the output signal would vary wildly according to Equation (5.3). Note that in this thought experiment, we have neglected the effect of a drifting effective signal frequency; in reality the phase drift will be random and nonlinear.

5.2.3 A Solution to the Phase Locking Problem

It is therefore an oversimplification to say that the phase offset of the reference signal must be matched with the phase offset of the signal (i.e. the phase at $t = 0$). If the signal frequency were constant but unknown, we could simply measure the frequency and phase offset of the signal once at $t = 0$, and then define the reference signal in Equation (5.5) with the constants $\omega_{ref} = \omega_{signal}$ and $\phi_{ref} = \phi_{signal}$. If the signal frequency actually continuously changes due to drift in f_{mod} or f_{sa} , then the cosine dependence of the output signal in Equation (5.3) applies to the total phase of the input and reference signals, not only the phase offset. Since the angular frequency is the rate of change of the phase, the signal will have a phase at time t of:

$$\Phi_{signal}(t) = \int_0^t \omega_{signal}(t)dt + \phi_{signal}. \quad (5.6)$$

The most intuitive solution would be to find the initial phase offset and the signal frequency at some sample times and estimate the total phase by numerically integrating over the frequency. This approach would probably be prone to error from the numerical integration and involve extra computation. In addition, because we have one condition (to match the reference

phase with signal phase) but two parameters (signal frequency and phase offset), it is convenient to cast the condition in terms of only one parameter, though this new parameter is less physically meaningful than the signal frequency and phase offset. We can satisfy the condition and avoid numerical integration by defining a constant reference frequency at the nominal setting and measuring an effective phase offset at all points $1 \leq n \leq N$. Satisfying this condition is represented by:

$$\Phi_{\text{signal}}(t[n]) = \int_0^{t[n]} \omega_{\text{signal}}(t') dt' + \phi_{\text{signal}} = \omega_{\text{mod}} t[n] + \phi_{\text{eff}}[n], \quad (5.7)$$

where for all n $\omega_{\text{ref}}[n] = \omega_{\text{mod}}$ and the reference phase offset is:

$$\phi_{\text{ref}}[n] = \phi_{\text{eff}}[n] = \int_0^{t[n]} \omega_{\text{signal}}(t') dt' + \phi_{\text{signal}} - \omega_{\text{mod}} t[n]. \quad (5.8)$$

This definition for the reference phase offset is an effective phase offset that can be interpreted as the phase difference of the measured signal from a sine wave with frequency ω_{mod} and zero phase offset. The definition is not useful as a formula to calculate it, since that would require measuring both signal parameters and performing the numerical integration. It is a statement of the existence of the effective phase offset and how it is related to the other more physical parameters. The remainder of this subsection is concerned with the algorithm to find the effective phase offset at all points as implemented at PEPL.

Finding the effective phase offset using phase-sensitive detection is computationally expensive, so the strategy is to find the effective phase offset at a sufficient number of points within the 60-s dataset so that the effective phase offset at all points $1 \leq n \leq N$ can be accurately interpolated between the set of points where it has been explicitly found. The number of points necessary can vary, but approximately ten to thirty points are often sufficient.

To find the effective phase offset at a point in the PMT signal, we apply phase-sensitive detection with $\omega_{\text{ref}}[n] = \omega_{\text{mod}}$ over a long integration time of about 100 ms to have a high signal-to-noise ratio in the PSD output signal ($\text{SNR} \gg 1$). Assume that the effective phase offset of the signal is approximately constant over the range of data where PSD is applied, approximately $t[n] - 2.5\tau$ to $t[n] + 2.5\tau$. This amounts to assuming that the signal and reference frequencies are sufficiently close that the phase drift over the integration time is small: $|f_{\text{signal}} - f_{\text{ref}}| 5\tau \ll 1$. This implies that $|f_{\text{signal}} - f_{\text{ref}}| \ll 2$ Hz for the typical $\tau = 100$ ms. This assumption is highly suspect with a modulation frequency on the order of megahertz and a frequency stability of many parts per million for the digitizer. The poorly satisfied assumption may explain the unreliable performance

and unexplained artifacts that appeared before the introduction of the external digitizer clock to improve frequency stability (described in the following subsection). Fortunately, the external digitizer clock improves the stability sufficiently to easily satisfy the condition, and the assumption of approximately constant effective phase offset is good. Now, if the effective phase offset is approximately constant during the integration time, then the PSD output signal will vary according to Equation (5.3) if the reference phase offset ϕ_{ref} is varied.

Phase-sensitive detection is applied several times with different reference phase offsets uniformly spread in the range $0 < \phi_{ref} < 2\pi$ (between about four and ten). The resulting signal varies as a $\cos(\phi_{ref} - \phi_{eff})$ function where the maximum occurs when $\phi_{ref} = \phi_{eff}$. The points can then be fit to a cosine function of the form $\cos(\phi_{ref} + \phi_{fit})$ using a nonlinear least squares fitting routine, where ϕ_{ref} is the independent variable and ϕ_{fit} is a fitting parameter. The effective phase offset is the reference phase offset at which the maximum occurs in the fit, or $\phi_{eff} = -\phi_{ref}$. An example of this procedure is shown in Figure 5.2. Equivalently but less intuitively, the points can be fit to a sine function of the form $\sin(\phi_{ref} + \phi_{fit})$, where in this case the effective phase offset is $\phi_{eff} = \pi/2 - \phi_{fit}$. The actual code uses the latter because a general sine function with fitting parameters appropriate for use with the MATLAB built-in nonlinear fitting function `lsqcurvefit()` was already available.

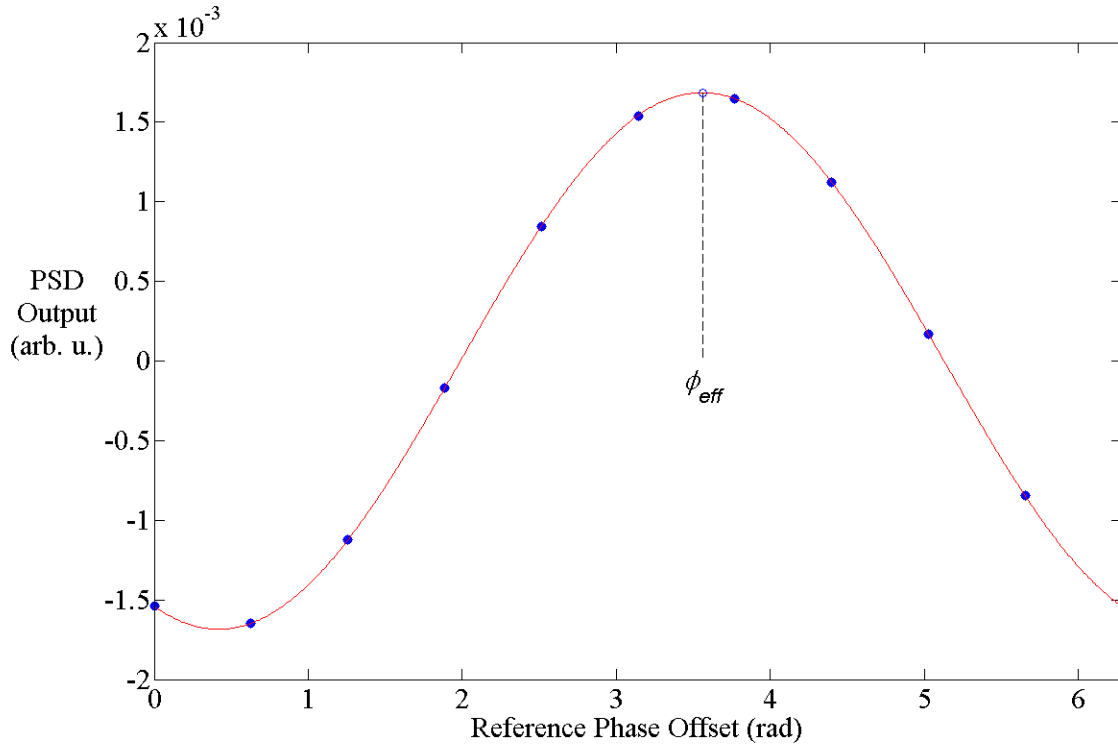


Figure 5.2. An Illustration showing the method of finding ϕ_{eff} at a single point in time. PSD output at ten different reference phase offsets (blue circles) are used in the least squares fit to the function $\sin(\phi_{ref} + \phi_{jit})$ (red line). The maximum occurs at $\phi_{ref} = \phi_{eff} = \pi/2 - \phi_{jit}$ (open blue circle).

Having found the effective phase offset at a uniformly spaced selection of points in the 60-s dataset, a function can be interpolated between them to find a reasonable estimate for the effective phase offset at all point within the dataset. In the best case, drift and noise are minimal, and a global linear fit with a relatively small number of measured points may be used. In the worst case, phase drift may be highly nonlinear and random and noise may be high, necessitating a larger number of measured points and a piecewise linear interpolation between small groups of neighboring points. Examples of these two extreme cases are shown in Figure 5.3. Each of the measured points in Figure 5.3 comes from a fit to a sinusoid such as in Figure 5.2. For example, the fit in Figure 5.2 produces the first point at $t = 1.5$ s in the top panel of Figure 5.3. The top panel comes from the H6 dataset in Chapter 7 with $I_{IM} = 1.8$ A at a laser wavelength of 834.94707 nm. The bottom panel comes from an early test of the TFLIF system with a cathode experiment similar to Chapter 6 from 15 September 2012.

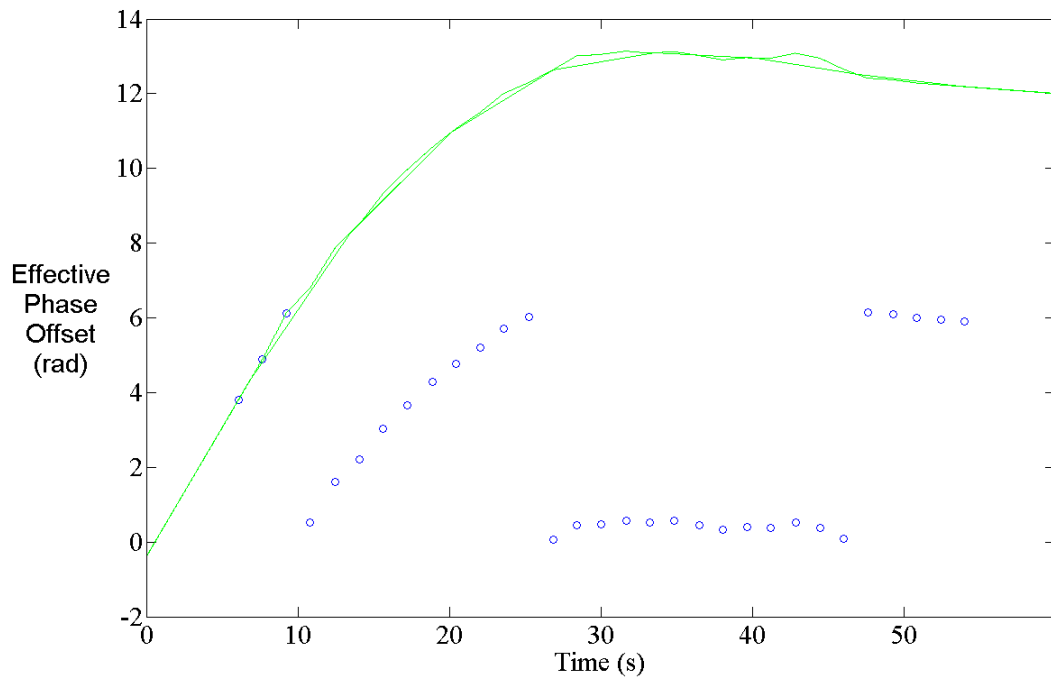
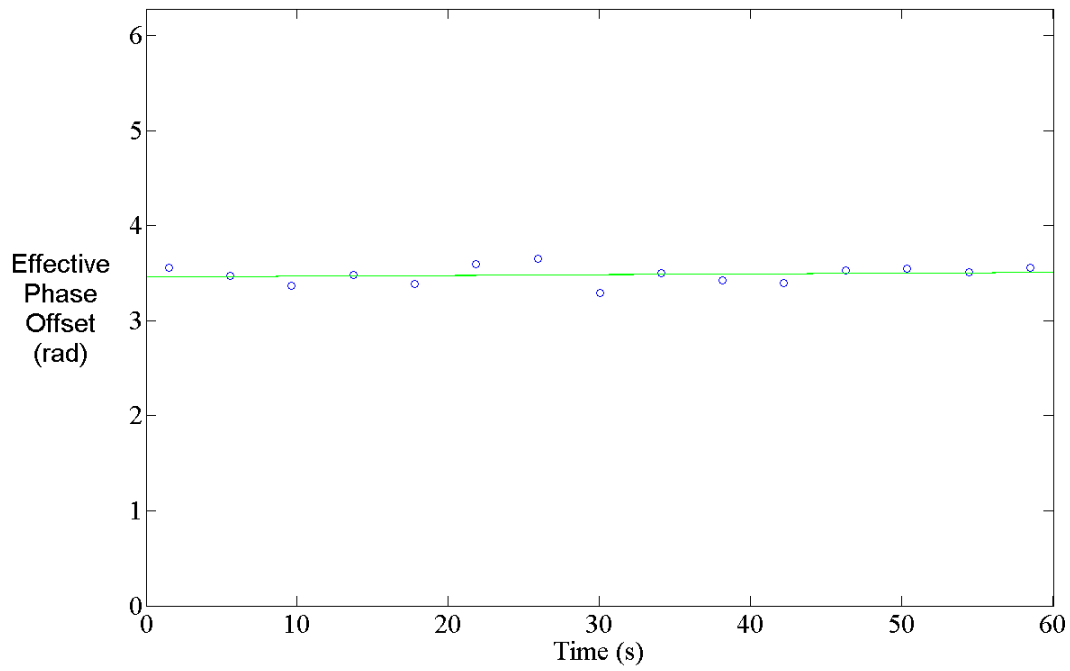


Figure 5.3. The effective phase offset at all points (green line) is interpolated from a series of effective phase offset points measured throughout the 60-s dataset (open blue circles). Top: example of the "best case" scenario with negligible drift and low noise. Bottom: A difficult example with fast, nonlinear drift in the effective phase offset.

5.2.4 External Versus Internal Digitizer Clock

In the discussion of frequency stability in the previous subsection, we found that the limiting factor was the internal sampling clock of the digitizer, having much worse nominal stability than any of the function generators considered. The Alazartech 9462 has options to use an external clock signal to time sampling. The "Fast External Clock" option is most appropriate for our purposes, accepting a sine wave with an amplitude between 200 mV and 1 V at an arbitrary frequency between 10 MHz and 180 MHz. Sampling can be selected on the rising or falling edge of the sine wave so that the sampling frequency is set by the sine frequency. Thus, the digitizer clock stability can be the frequency stability of the function generator providing the clock signal, down to 0.1 ppm for the Agilent 33522A instead of 25 ppm for the internal clock of the digitizer.

As shown in the top panel of Figure 5.3, an effective frequency stability much better than even 0.1 ppm can be achieved if the frequency stability of the digitizer clock is correlated with the AOM driving signal. Consider a digitizer clock signal that has the same relative frequency fluctuations as the AOM driving signal. A change in the AOM driving signal frequency by a factor of C will change the actual signal frequency by the same factor. But the sampling frequency will also change by the same factor, thus reducing the apparent frequency of signals by a factor of $1/C$, canceling the frequency fluctuation in the digitally sampled data. To see this effect more intuitively, consider the AOM driving signal itself used as the digitizer clock. Then one sample per modulation period will always be recorded regardless of fluctuations in the time base used to generate the signal. The AOM driving frequency ($f_{mod} \sim 1 \text{ MHz}$) is necessarily slower than the sampling frequency ($f_{sa} \sim 30 \text{ MHz}$), so the AOM driving signal cannot actually be used. The ideal solution is to use a two-channel function generator where the channels are synchronized by the same clock signal. Then the channels can generate distinct AOM driving and digitizer clock signals, but any frequency instability in f_{mod} is also in f_{sa} , in which case their effects cancel.

The Agilent 33522A is such a function generator. The "best case" with a nearly constant effective phase offset in Figure 5.3 above was from a dataset captured with this configuration. The "worst case" with badly nonlinear phase drift was from a dataset captured using the digitizer internal clock. Phase locking is still necessary but that small configuration change makes a dramatic difference in the phase drift observed. In hindsight, this configuration may seem

obvious given that the external clock option is available specifically for synchronizing the digitizer clock to some other signal for sensitive applications. As a new graduate student with no experience in signal processing and high-speed measurements, it took considerable time to identify the problems with phase locking and their solutions. Aside from the "worst case" example in Figure 5.3, all data presented in this dissertation were captured with the external clock using the Agilent 33522A to generate both the AOM driving signal and the digitizer clock. Results before this improvement were significantly noisier due to the difficulties in phase locking.

Prior to the use of an external digitizer clock synchronized with the AOM driver, the sampled frequency was often far enough from the nominal f_{mod} that the drift in the effective phase offset would be extremely fast. For example, the setting of the AOM is $f_{mod} = 500$ kHz for the bottom panel of Figure 5.3, but the sampled signal frequency is $f_{signal} = 500,005.4$ Hz. This leads to a fast phase drift of 33.9 rad/s between the reference and signal, or about 5 full periods per second. The fast drift is not visible in Figure 5.3 because the frequency has been corrected but this drift is shown in Figure 5.4 (note the time scale of 1 s). Phase drift that fast violates the assumption that the effective phase offset is approximately constant during the integration time on the order of 100 ms, and therefore the approach of Subsection 5.2.3 actually fails if applied directly in this case. The internal digitizer clock therefore requires an intermediate frequency fitting step to find the average signal frequency. A sample of the clean AOM driving signal from the function generator can be fit to a sine wave to determine the frequency with high precision. The reference frequency f_{ref} in the TFLIF software would then be set to this frequency (500,005.4 Hz) instead of the actual function generator setting ($f_{mod} = 500$ kHz). The intermediate frequency fitting step is unnecessary for the external clock configuration and in that case $f_{ref} = f_{mod}$.

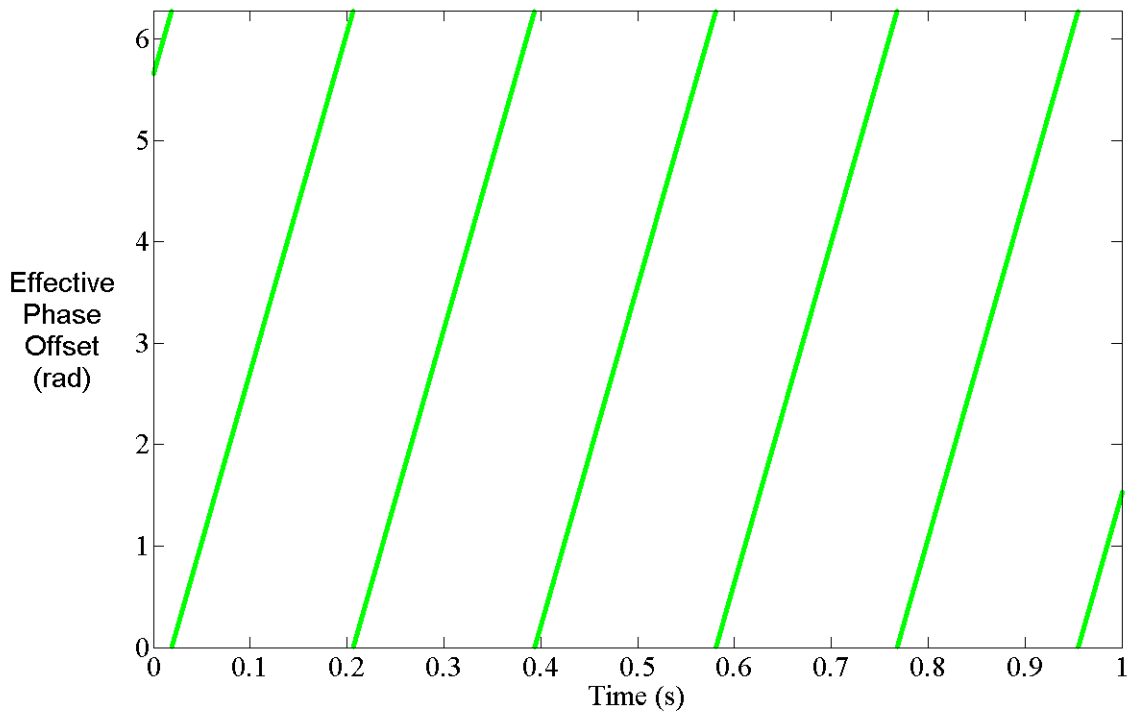


Figure 5.4. The internal digitizer clock can result in fast drift in the difference between the reference and signal phases (green line) that would break the algorithm in Subsection 5.2.3. An intermediate step fitting the modulation frequency is necessary in this case.

5.2.5 Phase Unwrapping

One further complication in phase locking made obvious by the bottom panel of Figure 5.3 is that the effective phase offset returned by the cosine fit is always between 0 and 2π . This can be a problem if phase drift or noise cause some points to cross over the boundary at 0 and 2π radians since those points will "wrap around" to the other side and a global linear fit will then be invalid. It is critical to detect the points that have crossed over the boundary from the others and unwrap them. Despite multiple $0/2\pi$ crossings in that example, unwrapping is relatively simple in such a low-noise situation. Crossings can easily be detected and accounted for by looking for a relatively large difference above some threshold between neighboring points.

Unwrapping becomes a more significant problem when noise makes the difference between neighboring points similar to the difference between points after a crossing, in which case a simple threshold may not correctly detect the crossings. As the SNR decreases, eventually the TFLIF signal is lost entirely because the effective phase offset fit fails, leading to erroneous

PSD output signals used in the transfer function average. Fine tuning the threshold can help but requires manually setting for each case and still may not correctly detect crossings in poor SNR.

5.3 Data Recycling

In calculating transfer function estimators, windowing the chunks of data before calculating the discrete Fourier transform can improve the final result (see Appendix A).

Consider the windowed discrete Fourier Transform:

$$\tilde{A}[k] = \sum_{n=0}^{N-1} w[n]A[n]\exp(-2\pi ikn / N), \quad (5.9)$$

where $w[n]$ is a discrete window function, which tend to be bell-shaped. Recall from Subsection 4.3.3 that the average transfer function is calculated by splitting the 60-s dataset into many chunks about 1 ms long and then averaging the transfer function estimators that are calculated for each chunk.

The most intuitive way to split the 60-s dataset into chunks is to split it into non-overlapping chunks like slicing a loaf of bread. One effect of the weighting of the window is that by definition points near the edge of each chunk do not contribute as much as points near the center to the resulting spectrum. When averaging many windowed spectra from the same continuous dataset together, as is done during the transfer function average, the result is that there are alternating sections of data where about 50% of the points have a relatively large contribution to the transfer function and the other 50% contribute almost nothing. The top panel of Figure 5.5 illustrates this idea by showing the window functions of three chunks plotted against the time axis. In addition to the non-uniform weighting of points, the discrete Fourier transforms of two adjacent chunks do not capture information about how points within one chunk are related to points within the other. Some information is clearly lost in this process.

This section describes a method of recovering more information from the same dataset by taking overlapping chunks instead of the basic non-overlapping chunks like slices of bread. Chunks are still taken sequentially, but the analysis code loops through multiple cycles of the dataset with a different offset for each cycle. The bottom panel of Figure 5.5 illustrates this concept for two cycles with the red chunks in the second cycle that overlap the blue ones from the first cycle.

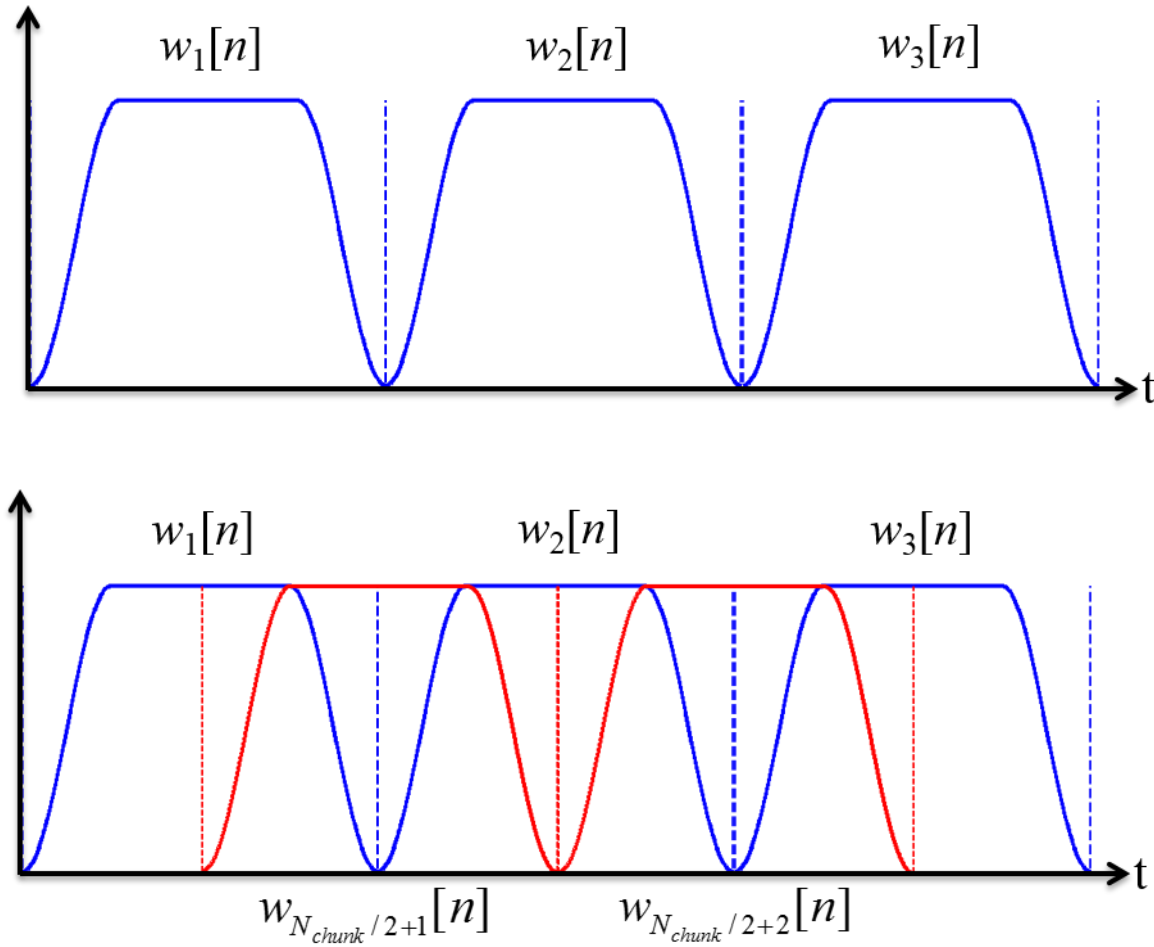


Figure 5.5. Top: A conceptual illustration showing how windowing consecutive chunks leads to large sections of data that do not contribute significantly to the DFT. Bottom: An example illustrating data recycling with $C = 2$, with the first cycle chunks shown in blue and the second cycle chunks shown in red. N_{chunks} is the total number of chunks taken from the dataset.

5.3.1 What is Data Recycling and Why is it Important?

One solution to recover more information from the same dataset is to calculate the average transfer function using chunks of data with overlapping windows instead of the simple picture of slicing the dataset like a loaf of bread into many independent chunks with no overlap. This idea was implemented independently at PEPL for transfer function averaging, but an analogous method for power spectrum estimation was later found in the literature. The overlapping averaging method is known as the Welch method [95] to distinguish it from the non-overlapping Bartlett method [96]. Note, the Welch method for periodogram averaging is a distinct idea from the Welch weighted transfer function average in Subsubsection 4.3.3.3.

The idea was called data recycling at PEPL since the same data points are "recycled" into different chunks in the average a number of times. The name data recycling is also apt since the implementation of it at PEPL has sequential chunks taken end-to-end from the temporary vector in memory (see Subsection 5.4.1) and then, after completing a full cycle through the data in memory, the analysis routine then takes another set of sequential chunks with some offset from the previous chunks. The bottom panel of Figure 5.5 illustrates this concept for two cycles; the chunks of the first cycle are in blue and the chunks of the second cycle are in red and offset from the first cycle by 1/2 of a window length.

A caveat of this overlapping averaging scheme is that the neighboring traces being averaged together become increasingly correlated as overlap increases, and therefore the average removes noise less efficiently. In the extreme case of only one point offset between chunks, the signal, noise, and therefore spectrum of one chunk will be virtually the same as its neighbors, and therefore the extra chunks will not reduce noise. An analogy is measuring a constant voltage in the presence of noise, where averaging N independent measurements would improve the SNR by a factor of \sqrt{N} , but averaging two copies of the N independent measurements together would make no further improvement, of course.

Considering that calculating the empirical transfer function has the same computational expense for all chunks whether independent or not, the most computationally efficient average (i.e. the best ratio of SNR improvement to computational expense) would be to take completely independent chunks. To average a similar number of chunks to gain the same SNR improvement implies a longer acquisition time and more difficult storage requirements. Thus, ultimately data recycling represents a tradeoff where some computational efficiency is sacrificed to reduce both data acquisition time and storage requirements. The tradeoff can be tuned somewhat to the needs of the experiment by adjusting the overlap, or the number of cycles.

5.3.2 A Test of Data Recycling

The key question to help determine the optimal tradeoff is: how much overlap can be used, or equivalently how many cycles can be used, in the average while still acceptably improving SNR? This is a complicated question to answer analytically, but we can attempt to answer this question empirically by testing the effect of varying the number of cycles used in the average.

We are actually interested in the error of the TFLIF signal, or the difference between the TFLIF signal and the exact TRLIF signal without noise, but the exact signal is unknown. Instead, the easiest metric in this case is to compare the TFLIF signal from C cycles of analysis with the triggered average signal. The periodic cathode oscillation dataset from the validation test of Section 6.3 is used for this test. In the case of periodic oscillations, both the TFLIF signal and TALIF signal converge to the exact TRLIF signal (see Subsubsection 4.3.3.4 and Subsection 4.3.4). The residual between the TFLIF signal and TALIF signal is the difference between the errors of both measurements. In the formalism introduced in Chapter 4, the residual is:

$$\begin{aligned}
R &= F_{characteristic}[n] - \langle F_q[n] \rangle \\
&\approx (F_{av}[n] + N_{characteristic}[n]) - (F_{av}[n] + N_{TA}[n]) \\
&= N_{characteristic}[n] - N_{TA}[n].
\end{aligned} \tag{5.10}$$

Thus, this metric can be used to track changes in $N_{characteristic}[n]$ since the error in the TALIF signal is constant. Since the triggered average is generally more efficient than the transfer function average, for a small number of cycles the error in the transfer function average will dominate the residual and $R \approx N_{characteristic}[n]$. The potential downside of this metric is that the RMS residual will approach a constant as $N_{characteristic}[n]$ approaches zero since the constant $N_{TA}[n]$ will dominate. Thus, the RMS residual is an upper bound on the RMS of $N_{characteristic}[n]$ that converges to the RMS of $N_{characteristic}[n]$ if $\text{RMS}(N_{characteristic}[n]) \gg \text{RMS}(N_{TA}[n])$.

Unlike all other data in this dissertation, the dataset in this section was analyzed with the empirical transfer function estimator (ETFTE), not the Welch weighted estimator (see Subsubsections 4.3.3.2 and 4.3.3.3), in order to maximize noise in the transfer function average to ensure that it dominates, which is confirmed by inspection of Figure 5.7, except that noise appears to be comparable with the noise in the TALIF signal by about $C = 10$. There is a clear progression of noise reduction as the number of cycles C increases. Note that analysis was done with a Tukey window, and results may vary somewhat depending on the window used. The Tukey window is the standard used in most TFLIF analysis in this dissertation.

Figure 5.6 presents the results of this test in terms of the RMS residual for a number of cycles between one and ten. The measured RMS residual (blue diamonds) falls increasingly short of the A/\sqrt{C} trend that would be expected for averaging over independent chunks (red squares) as C increases. The measured RMS residual for $C = 2$ is within 10% of the expected value for

independent chunks because the chunks are indeed largely independent of each other due to window weighting. For all other $C > 2$, the measured RMS residual is significantly larger than that expected for averaging independent chunks, though there is still significant improvement all the way up to $C = 10$.

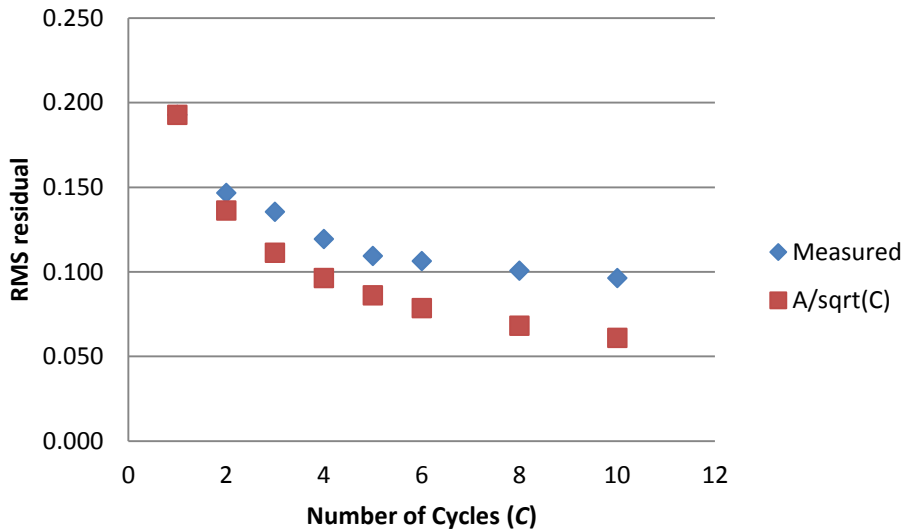


Figure 5.6 The measured RMS residual between TFLIF and TALIF signals as a function of the number of cycles in the transfer function average (blue diamond). Also shown is the expected RMS residual following the A/\sqrt{C} function expected for averaging over the same number of chunks if they were independent. The measured noise reduction closely follows the independent averaging trend only for $C = 2$, and lags increasingly further behind thereafter.

Most experiments would certainly benefit from at least using $C = 2$, since it can provide almost the same SNR improvement as averaging over independent chunks without actually measuring and storing twice as much data, and do so with the same computational expense. Evidently from Figure 5.6, $C = 5$ provides at least the same noise reduction as averaging over independent chunks in a dataset 3 times the size. Put another way, data recycling allows the acquisition time and data storage to be reduced by a factor of 3 while increasing computation time by about 5/3. Similarly, $C = 10$ provides at least the same noise reduction as using only independent chunks in a dataset 4 times as long. Therefore, to achieve the same SNR, data recycling allows the acquisition time and data storage to be reduced by a factor of four while increasing the computation time by about a factor of 2.5.

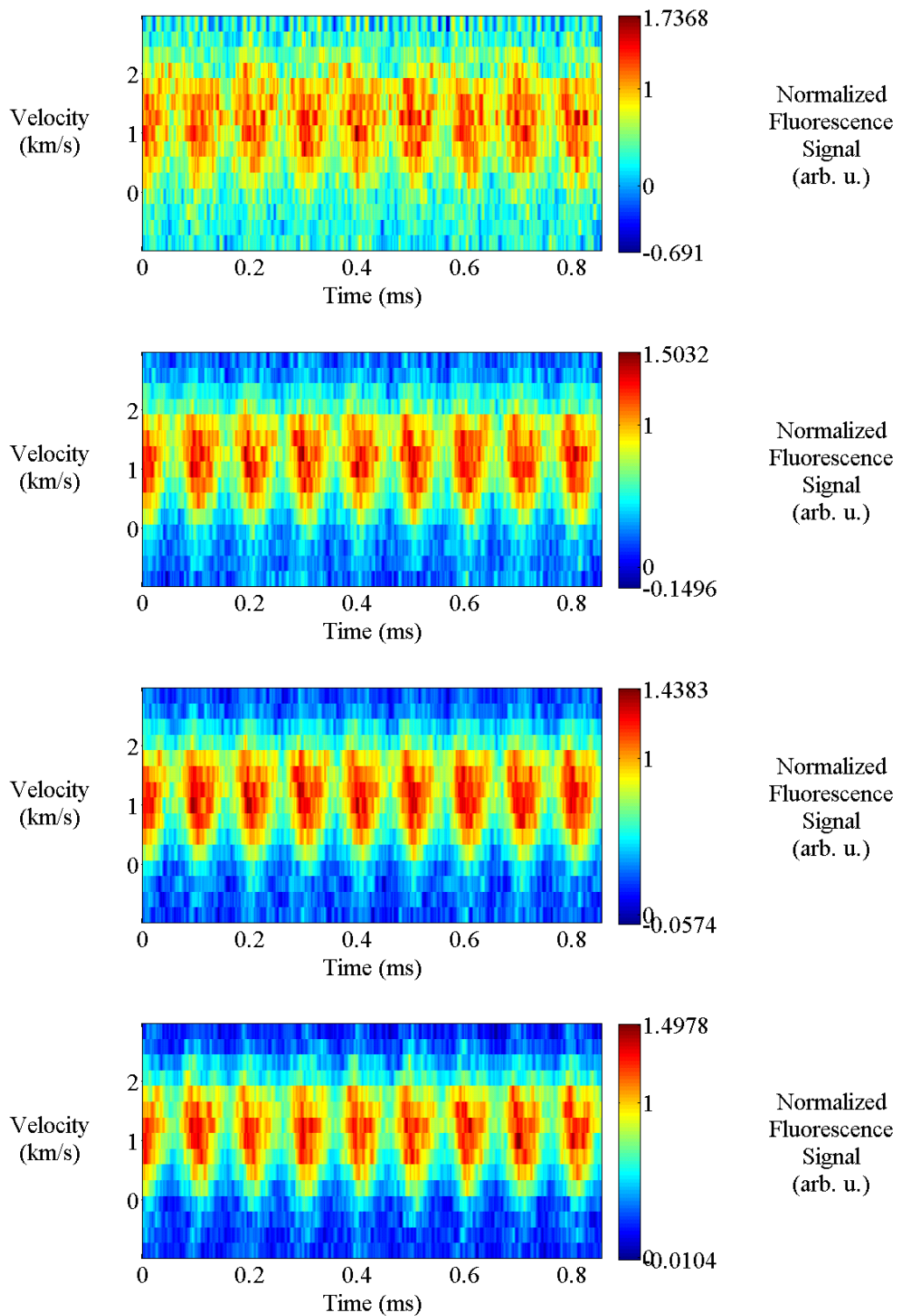


Figure 5.7. TFLIF signal for $C = 1$ (top), $C = 5$ (upper middle), and $C = 10$ (lower middle) converge toward the TALIF signal (bottom) as the number of cycles C increases.

5.3.3 Guideline for Setting the Value of C

Consider a standard dataset for a single wavelength that is composed of three channels (one output and two inputs) of 16-bit samples with a sample speed of 30 MHz for a duration of 60 s. These are the parameters of the most recent experiment, the demonstration of TFLIF on a Hall thruster in Chapter 7. All other datasets in the dissertation are similar, though previous ones only captured two channels (one input only). These three-channel datasets are approximately 10 GB, leading to 200 GB of raw data for a complete VDF with 20 wavelengths. The sheer size of these datasets is cause for concern, even for the 14.5 TB RAID, so it is clear that recovering as much information as possible would be important.

Determining the optimal number of cycles may be subjective, but we can recommend a reasonable guideline based on an interpretation of the results from the previous subsection and the storage concerns above. In addition to reducing the data storage requirements, the reduced acquisition time from data recycling allows many more points, reduces the risk of laser and thruster issues occurring during the experiment, and reduces the expense of xenon and other facility costs incurred while operating for a longer time. The relatively modest increase in computation time is a very attractive tradeoff to gain all of that. In fact, since the test shows significant improvement even up to $C = 10$, a general guideline may be to simply use the largest C possible given the time allowed for computation.

Each of the 20 wavelengths for a complete VDF took a duration of about 1 hour and 45 minutes to analyze. Even with parallelization according to Subsection 5.4.2, a single spatial point takes the majority of a day in computation time. Therefore, a large number of points and possibly a variety of analysis parameters could easily take weeks of analysis. A significantly longer computation time would not be tenable and will have diminishing returns in any case. In light of these considerations, $C = 10$ is the standard by which all datasets outside of this section are analyzed, and it is probably a reasonable choice for the needs of most experiments.

5.4 Other Software Considerations

The data directly measured by the TFLIF system, the band-pass filtered voltage drop across the terminating resistor and the discharge current, are very raw compared with the directly measured and saved data of other TRLIF techniques, which typically only save the resulting time-resolved waveforms after processing by the instruments. In the case of TFLIF, all analysis on the raw data is done in post processing after the experiment. This leads not only to the

hardware considerations in Section 5.1 on data transfer and storage, but it also leads to challenges in the software implementing the analysis described in Chapter 4. This method does have the advantage, however, of the ability to perform analysis many times to fine tune parameters and even apply completely different analysis techniques to the same dataset for comparison, which is used to our advantage as a key part of validation.

The goal of this section is to introduce a few of the considerations encountered in implementing TFLIF in software. The first two subsections describe at a high level the memory management (Subsection 5.4.1) and parallelization (Subsection 5.4.2) of the analysis code without encumbering the reader in the minutia of the actual code or even pseudocode. Subsection 5.4.3 shows why it is best to limit the chunk length to certain numbers of points. Finally, Subsection 5.4.4 gives an example of one advantage of having most of the data analysis in post-processing: optimal analysis parameters, such as the PSD time constant, can be found by testing a selection of different values.

5.4.1 Memory Management and Temporary Vectors

If datasets are large enough that simply transferring and storing the raw data are a significant issue (see Subsection 5.1.2), then surely having enough relevant data in system memory for the analysis is also a consideration. The TFLIF PC currently has 72 GiB of RAM, incrementally upgraded as needed from the 12 GiB originally installed in the PC. A 60-s, 3-channel dataset captured at a sample frequency of 30 MHz is approximately 10 GiB, but that is stored in 16-bit integer format. It is most practical and accurate to do calculations on double-precision floating point numbers, which occupy 8 bytes of memory, or four times as much memory as 16-bit integers. Hence, simply reading a complete dataset for a single laser wavelength into memory would require more than half of the spacious system memory. In addition, the memory occupied by other variables used in the computations typically require several gigabytes or tens of gigabytes more memory than the raw dataset.

The system memory could handle these requirements for a single dataset, but all of the data for analyzing several datasets must be kept in memory simultaneously due to the parallelization discussed in Subsection 5.4.2. The only solution in this case, barring somehow upgrading the already spacious system memory by about an order of magnitude (not possible for this main board), is to keep not all of each dataset in system memory simultaneously. Since calculations are done on one chunk of about 1 ms at a time to calculate a single transfer function

estimator, it is not necessary to read all of the raw dataset at once. On the other hand, it would be highly inefficient to read only one chunk of raw data at a time. The primary reason is simply that reading from permanent storage is the largest bottleneck for a PC, even for the RAID. The maximum data rate of the (DDR3 1333) RAM is 10666.67 MB/s, about an order of magnitude faster than the RAID. Due to data recycling, if data were only read into RAM one chunk at a time, then the same raw data would be read from permanent storage as many as C times in one analysis, where C , the number of cycles, is about ten.

The balance between those two considerations used in the TFLIF code is to transfer from storage to RAM a fraction of the complete dataset that is much more than a single chunk, but much less than the complete dataset. This fraction of the dataset that is read into memory at one time is called a "temporary vector." It is temporary in that it is only in memory for part of the analysis time and is replaced by the next temporary vector when the code is finished calculating transfer function estimators from all of the possible chunks in the temporary vector. It is a "vector" in the sense of the term in MATLAB: it is a 1D array of data. In this sense, there is a temporary vector containing the time series of each of the digitizer channels used in analysis.

The temporary vectors of 16-bit integers are read from permanent storage to RAM. The length of the temporary vector in memory is a number of points that is a power of two and as close as possible but without being larger than some limit, usually set at 3 GiB. In this case, there are a few temporary vectors per dataset. The chunks are read from the temporary vector in sequence and converted to double precision and scaled according to the digitizer settings before calculating transfer function estimators. The data recycling of Section 5.3 actually occurs over each temporary vector, not over the complete dataset since the ends of both temporary vectors would be needed to read chunks that overlap the ends between two temporary vectors. A workaround is of course possible, but would add extra complexity for a negligible number of chunks, so for simplicity those possible chunks overlapping the ends of the temporary vectors are lost.

5.4.2 Parallelization

Knowing in advance that the analysis would be extremely computationally expensive, the TFLIF PC was built with two physical Xeon X5650 processors, each with six physical cores, for a total of 24 threads with Hyper-Threading. In general, a single MATLAB instance will only load one or a few threads when using some of the in-built functions that are already parallelized.

Computing the average transfer function for each single-wavelength dataset takes between one and two hours for a single, serial process. Considering that there may be about 20 wavelengths per VDF, it would take well over a full day to generate the time-resolved VDF profile at a single point in space for a single set of analysis parameters. The total processing time could quickly become impractical for many VDF points at multiple positions or thruster operating conditions with analysis done several times with different parameters. Hence, it is critical to parallelize the analysis to utilize all of the available computational power.

The most obvious parallelization option in MATLAB is the Parallel Computing Toolbox containing many built-in functions designed to make use of multiple cores. For example, the `parfor` function acts as a for loop that processes the iterations of the loop in parallel, if the iterations are independent of each other, of course. This could be useful, for example, to compute transfer function estimators for many different chunks simultaneously.

Unfortunately, this approach has a number of limitations and complications. There is an artificial limit imposed by MATLAB on the number of "workers" available, i.e. the number of threads that MATLAB is capable of utilizing, or the number of transfer function estimators it may simultaneously calculate. This was limited to 8 workers when first writing TFLIF code, but later expanded to 12 in MATLAB R2011b. The limit can be raised by purchasing a MATLAB Distributed Computing Server (MDCS) license. In addition, it is difficult to work with and debug, with cryptic errors and arcane syntax and requirements for variables within the `parfor` loop. There is considerable overhead in the client managing the workers. There are five different classifications of variables within a `parfor` loop, and one must be careful to use them properly. For example, there is an option to use "sliced variables," where each worker works on independent segments of a variable (the temporary vector, for instance), but that introduces complications with data recycling. There is also considerable overhead and wasted memory in creating copies of variables that cannot be simply sliced for use in each worker or broadcast from the client without the workers changing them. The overhead was so expensive that, with all 8 workers active, the parallelization was found to reduce the computation time by only about a factor of two.

There are also third party packages available with varying degrees of flexibility and usability. MATLAB built-in `parfor` code was abandoned and the TFLIF code was parallelized using Multicore package by Markus Buehren. It uses multiple instances of MATLAB as workers

that communicate with each other through access to a directory. The main advantage over the built-in parallelization is that any number of workers can be used by opening instances of MATLAB, and it can be parallelized across many machines that have access to a common directory. Unfortunately, this package also has some syntactic limitations and overhead issues, which are impossible to avoid for any code that implements true parallelization of a single task. It was also abandoned due to disappointing improvement in the total computation time.

The above attempts focused on parallelizing the analysis of a single-wavelength dataset in the loop of computing transfer function estimators. There is, in fact, no need for that because the total analysis involves completely independent calculations of the average transfer function estimators for each wavelength. The Multicore package demonstrated that many instances of MATLAB can run simultaneously without errors or conflicts, though sometimes an instance may inexplicably crash when a large number of instances (> 10) are running.

Then much of the overhead can be eliminated by making the analysis embarrassingly parallel. This is the scheme currently in use. Many independent instances of MATLAB are set to operate on different single-wavelength datasets and therefore have no overhead in communication between workers, creating extra copies of variables, etc. Each worker simply executes the un-parallelized code for a single dataset at a time. The computation time for a single dataset is about the same or a little worse due to the increased resource utilization, but the total computation time for a complete VDF is nearly inversely proportional to the number of workers. Quickly finishing the analysis on a single dataset is seldom desired, so there is little motivation to truly parallelize the analysis on the level of a single dataset.

This scheme works but becomes tedious to open and manually set a separate instance of MATLAB for each single-wavelength dataset. It is usually only practical to run about 10 instances of MATLAB on the TFLIF PC, so the workers would also have to be restarted on a second dataset to complete the 20-wavelength VDF profile. A simple but powerful semaphore scheme was developed that allows the individual MATLAB instances to determine which dataset to analyze based on which ones are already done or in progress. When a worker finishes with a dataset, it selects another appropriate one to analyze until there are no unanalyzed datasets left.

5.4.3 Chunk Length and the Power of a Power of Two

We conclude the section of software considerations with an example of a detail that is not at all obvious to someone unfamiliar with the details of the FFT algorithm but that makes a

dramatic difference nonetheless. The Fast Fourier Transform (FFT) is the class of algorithms most commonly used to calculate the discrete Fourier transform (DFT). Applying the definition of the DFT in Equation (4.43) directly to a time series of length N results in $O(N^2)$ operations. Fast Fourier Transform algorithms reduce the number of operations to $O(N \log_2 N)$, resulting in orders of magnitude in computational savings for large N . The actual number of operations, and therefore the computation time, is highly sensitive to the length N and can vary widely even for very small variation in N down to just one point.

This is relevant for TFLIF because the FFT is used to calculate the discrete Fourier transforms as described in Subsection 4.3.3. A transfer function estimator for each chunk requires an FFT performed on both the output and input time series. For example, for the H6 data set using a chunk length of $2^{14} = 16384$ points (about 5.5×10^{-4} s at a 30 MHz sampling frequency), the analysis averaged a total of slightly over one million chunks (using data recycling with $C = 10$). This means that the FFT was done over two million times per average transfer function estimator, taking a significant fraction of the analysis time. For that length, the FFT uses only about one ten-thousandth of the number of operations required to apply the DFT definition directly.

The details of the FFT algorithms are beyond the scope of this dissertation. It suffices to say that they save computation by splitting the time series of length N into smaller subseries for which the DFT is easier to calculate. The full DFT of length N is calculated from the shorter ones by exploiting certain mathematical properties. The first FFT algorithms assumed that the length N of the large time series was a power of two so that it could be evenly split into $\log_2 N$ subseries. The Cooley-Tukey algorithm that popularized the FFT is more general in that it applies to calculating the DFT of a time series with a composite (non-prime) length [97].

The actual effect of time series length on the calculation time of the MATLAB FFT function was investigated systematically for every length between 2^{14} and 2^{15} (corresponding to actual chunk lengths between 5.5×10^{-4} s and 1.1×10^{-3} s for the 30-MHz sampling frequency). For each time-series length, 600 samples of random time-series were generated and the time to calculate the FFT was measured with little other background activity on the PC. The standard deviation estimator of the 600 samples was typically large, often on the order of the average time itself. This suggests that the actual calculation time is highly influenced by whatever bottlenecks

are randomly encountered by the PC at the time rather than being strictly proportional to the number of floating point operations.

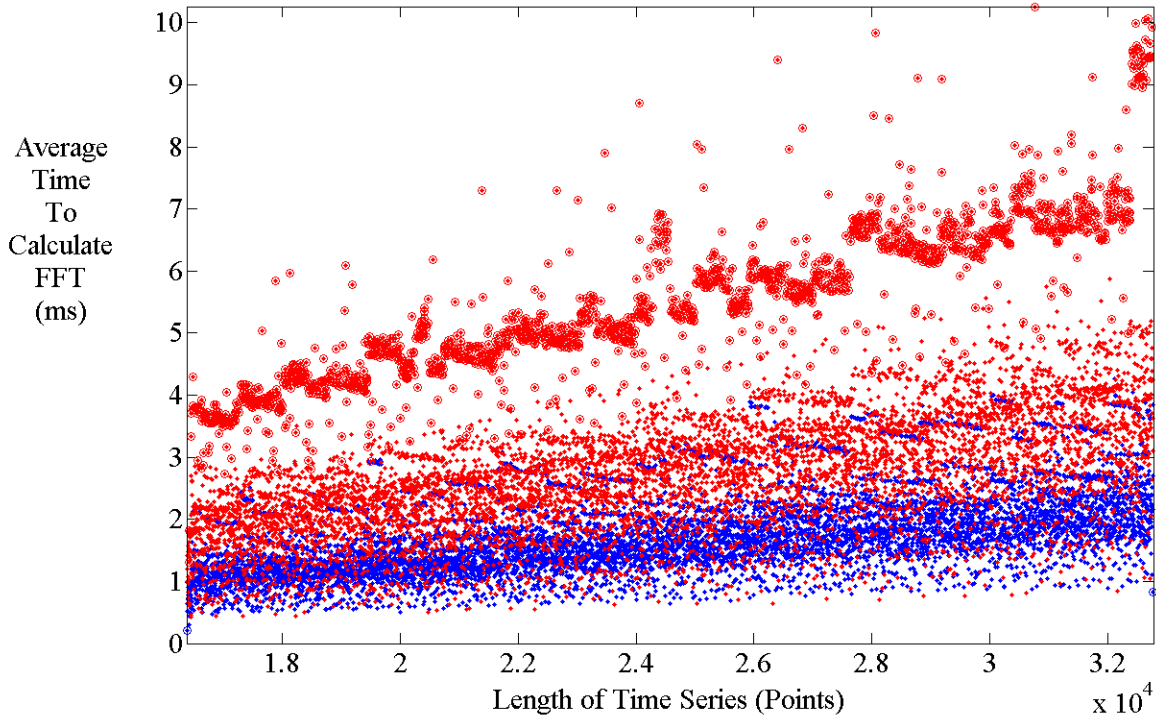


Figure 5.8. Plot of the average computational time of the FFT for all time series lengths between 2^{14} and 2^{15} . Even lengths (blue dot) are generally faster than odd lengths (red dot). Moreover, prime lengths (red circle) perform worst and powers of two (blue circle) perform best.

Even so, clear trends emerge in the plot of average FFT calculation time in Figure 5.8. The results vary apparently randomly and with some regular structure beyond the simple $O(N \log N)$ scaling. This test shows that the powers of two (the endpoints of 2^{14} and 2^{15}) significantly outperform all comparable lengths. Lengths that are even (blue) generally fare better than odd lengths (red). Generally, we see that it is possible for non-power-of-two and odd lengths to perform nearly as well as powers of two, but it is much more likely that the average calculation time will be at least about a factor of two higher; even up to an order of magnitude more for some odd lengths.

The consequence for TFLIF is that chunks are always chosen to be a power of two; otherwise the analysis could take significantly longer to complete. The total analysis time would not change as dramatically as Figure 5.8 suggests because the total analysis includes significant parts outside of the FFT that are invariant to chunk length, e.g. reading temporary vectors and

phase locking. Restricting the chunk length to a power of two means that round numbers for the total time of a chunk are generally not available. While the chunk length may be identified as "about 1 ms," the actual time spanned by a chunk is closer to $(2^{15}-1) / 30 \text{ MHz} \approx 1.09 \text{ ms}$.

5.4.4 Setting the Time Constant of Phase-Sensitive Detection

Equation (4.34) provides a double inequality condition on the time constant that ideally should be satisfied, but in practice may be only a crude approximation. Equation (4.39) provides more precise bounds on the time constant, but the maximum relevant frequency component f_{max} may not be easily defined. As discussed in Section 4.3.2, increasing the time constant of phase-sensitive improves noise rejection by reducing the bandwidth that the low-pass filter passes. Since PSD applies a moving, weighted average over a length of time about 5τ , increasing the time constant also destroys time resolution. In practice, the optimal time constant should be as large as possible to maximize noise rejection while ensuring that the filter bandwidth is sufficient to produce at most negligible distortion in the waveform.

One of the advantages of acquiring data in a raw state for post-processing is that it enables analysis of the same dataset many times with different parameters or analysis techniques. The optimal values for the analysis parameters set in software need not be known in advance because this feature allows them to be determined after the experiment.

Then a straightforward, if computationally intensive, way to determine the optimal setting for τ is to test a selection of values. An example of this test for the H6 at the 150-V, 10-mg/s operating condition with nominal magnetic field at the upstream axial position of $z = -4 \text{ mm}$ is shown in Figure 5.9. There is a distinct visual change in the noise level but little noticeable waveform distortion in the heat maps.

It is important to note that the time constant of the low-pass filter in phase-sensitive detection is adjusted. This is at the step of demodulating the signal prior to transfer function averaging, not simply applying a low-pass filter to the final resulting TFLIF signal. The net effect is similar because PSD passes a low-pass filtered version of the signal envelope [90], which is then used in the transfer function average. The low-pass filtering effect is then built into the average transfer function.

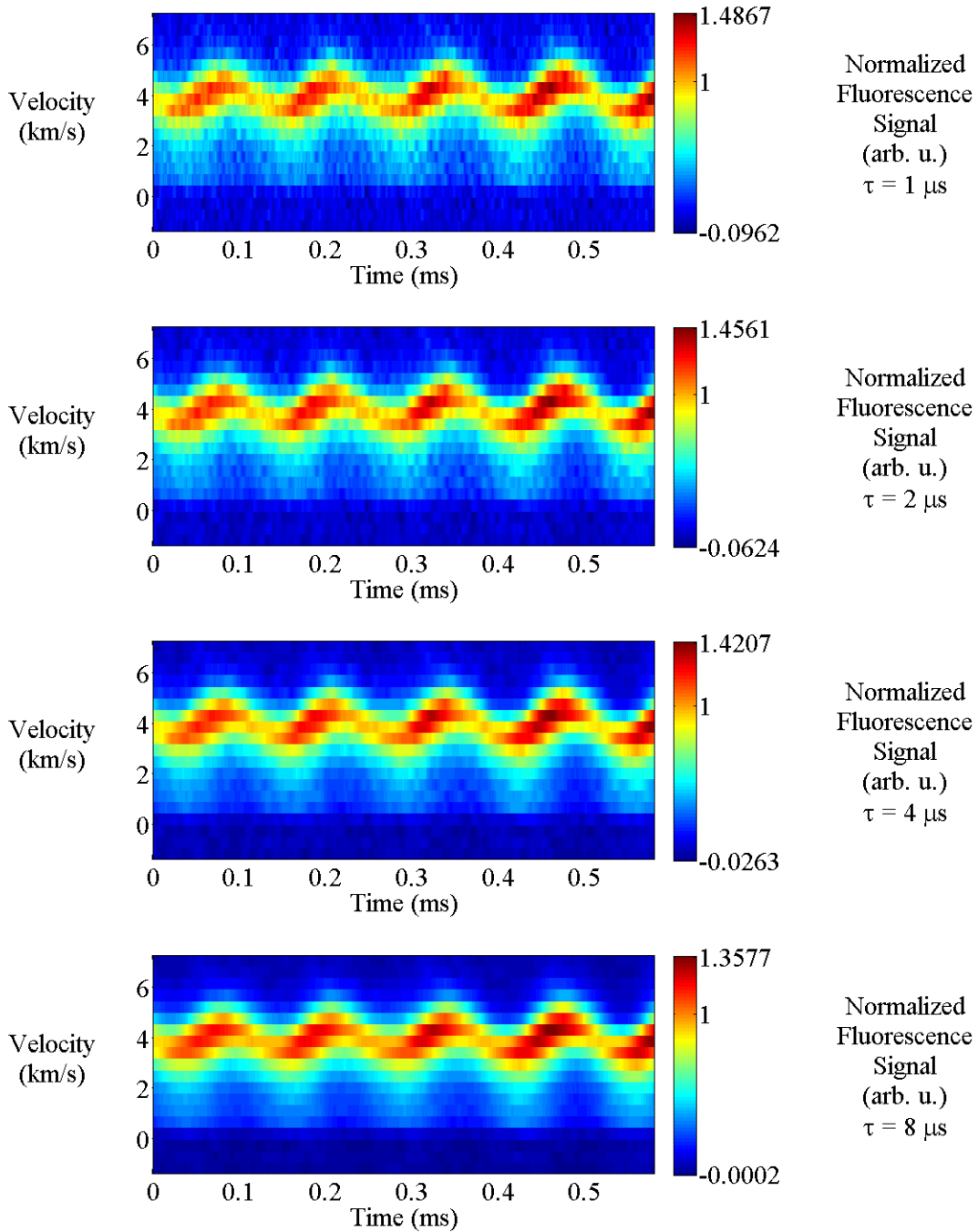


Figure 5.9. Raising the time constant from $1 \mu\text{s}$ to $8 \mu\text{s}$ markedly improves SNR while producing barely noticeable distortion of the waveform.

It is difficult to see distortion in Figure 5.9 by eye. Minor distortion becomes visible for the higher time constants when the waveform with $\tau = 1 \mu\text{s}$ is subtracted from the waveform with a higher time constant to reveal systematic differences. In Figure 5.10, faint blue streaks begin to appear for $\tau = 4 \mu\text{s}$, and distinct (but still small) patches of systematic error from the $\tau = 1 \mu\text{s}$

waveform are present in the $\tau = 8 \mu\text{s}$ waveform. For any given time series corresponding to a single velocity group, the peak is reduced in intensity and the wings are broadened in time. For the time-resolved VDF as a whole, the minimum instantaneous mean velocity is higher and the maximum instantaneous mean velocity is lower by about 100 m/s for the $8 \mu\text{s}$ case.

Though some minor distortion was detected, none of the time constants tested produced an unacceptable distortion. The time constant $\tau = 2 \mu\text{s}$ was conservatively chosen as the standard time constant for the results shown in the remainder of this dissertation because it is the largest time constant that produces no distortion above the noise floor. It is important to emphasize that this time constant setting is not a general result and could change if the signal bandwidth changes, for example.

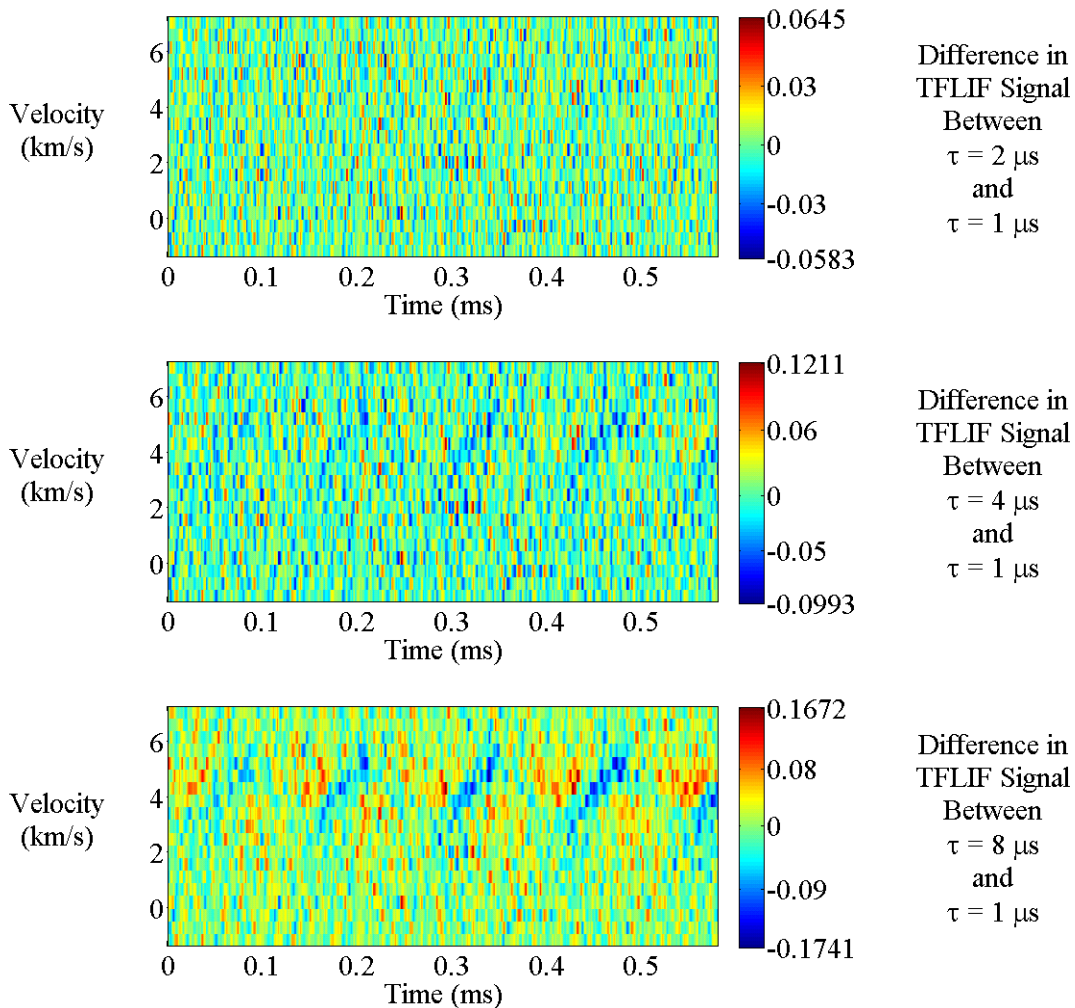


Figure 5.10. Plots of the difference in TFLIF signal waveform analyzed with $\tau = 2, 4,$ and $8 \mu\text{s}$ from the waveform with $\tau = 1 \mu\text{s}$. The systematic differences for $\tau = 4 \mu\text{s}$ and $\tau = 8 \mu\text{s}$ imply waveform distortion, while distortion present for $\tau = 2 \mu\text{s}$ is below the noise.

5.5 Optimizing the Signal-to-Noise Ratio

Using the highest laser modulation frequency possible is generally best due to the reduced noise spectral density at higher frequencies (Subsection 5.5.1) and improved time resolution (Subsection 4.3.2). Unfortunately, it was discovered early on that the signal-to-noise ratio of the TFLIF or time-averaged LIF measurement reaches a maximum at some modulation frequency and falls monotonically thereafter. One important part of the development of the TFLIF system has been to identify and resolve various limiting factors that cause the loss of signal at high modulation frequencies. Another important part of the development is to determine the optimal settings for parameters that maximize the SNR. This section is focused on finding and resolving those limits to the modulation frequency as well as finding those optimal parameters that maximize the SNR.

In the beginning, the maximum in SNR occurred near 100 kHz, a modulation frequency that would make Equation (4.39) difficult to satisfy. The first limits identified and solved were relatively simple issues related to the implementation of the technique. Following a suggestion from Professor Azer Yalin at Colorado State University, it was found that the combination of the internal capacitance of the photomultiplier and cabling and the resistance of the terminating resistor produced a significant RC constant that low-pass filtered the signal. The original 90 k Ω resistor was replaced with a 1 k Ω resistor to ensure the low-pass cutoff frequency is well above the working modulation frequency. A large resistor was chosen in the first place to amplify the voltage drop across the terminating resistor to a level measureable by the digitizer. Larger input and output gain settings of the Krohn-Hite 3945 are now used to compensate for the reduced voltage. Another early limit was the phase-locking discussed in Section 5.2. Higher modulation frequencies were more difficult to phase lock before the use of the external digitizer clock due to severe effective phase drift. The remainder of this section covers the limits to maximum of the SNR as a function of modulation frequency that have been encountered after early issues such as these.

Measurements of the noise spectral density at a few spatial locations and operating conditions of the H6 are given in Subsection 5.5.1. Subsection 5.5.2 quantitatively estimates the raw SNR of LIF data to show why averaging over a large quantity of data is necessary. Subsection 5.5.3 explains how the SNR is quantified to determine the highest usable modulation frequency. Subsection 5.5.4 provides an early test of SNR vs. modulation frequency with the

maximum at 500 kHz. Subsection 5.5.5 shows that the AOM can be a limiting factor and that the sampling frequency is not a limiting factor at the current operating parameters. Subsection 5.5.6 shows that the maximum occurs at about 2 MHz for the H6, further evidence that the modulation speed limit depends on the properties of the plasma source. Finally, Subsection 5.5.7 shows that the final SNR is nearly invariant to the band-pass filter setting, confirming the prediction of Subsections 4.3.1 and 4.3.2.

5.5.1 Noise Spectral Density

Consider the noise in the digitized photomultiplier signal. We consider “noise” to be any measured signal that obscures the desired LIF signal, including background signal (non-LIF collected light), fundamental noise sources (e.g. shot noise and Johnson noise), and other various sources of noise such as amplifier input noise and interference pick-up from external sources. This definition of noise is convenient since all of these sources can influence the SNR of the measurement in a complicated way depending on many different variables. The total noise spectral density can be measured by analyzing the photomultiplier signal with the laser off. Though this approach excludes any noise sources fundamentally linked to the signal itself such as shot noise and quantization noise from the digitizer, we will see that the overwhelmingly dominant portion of noise is the background light collected by the optics that was emitted by the plasma. The noise in the PMT signal is measured under the same conditions as the LIF signal, with same setting for PMT bias and monochromator pass band. The monochromator is set to a pass band of about 1 nm, so background emission is collected only near the transition wavelength.

The linear noise spectral density was calculated from the potential drop across the PMT terminating resistor captured by the digitizer. The calculation used the Welch technique averaging over 128 modified periodograms with zero overlap using a Hann window [95]. The effect of the equivalent noise bandwidth of the window was taken into account [98] and a consistency check verified that the square root of the integral of the power spectral density (not shown) is equal to the RMS observed in the time domain signal.

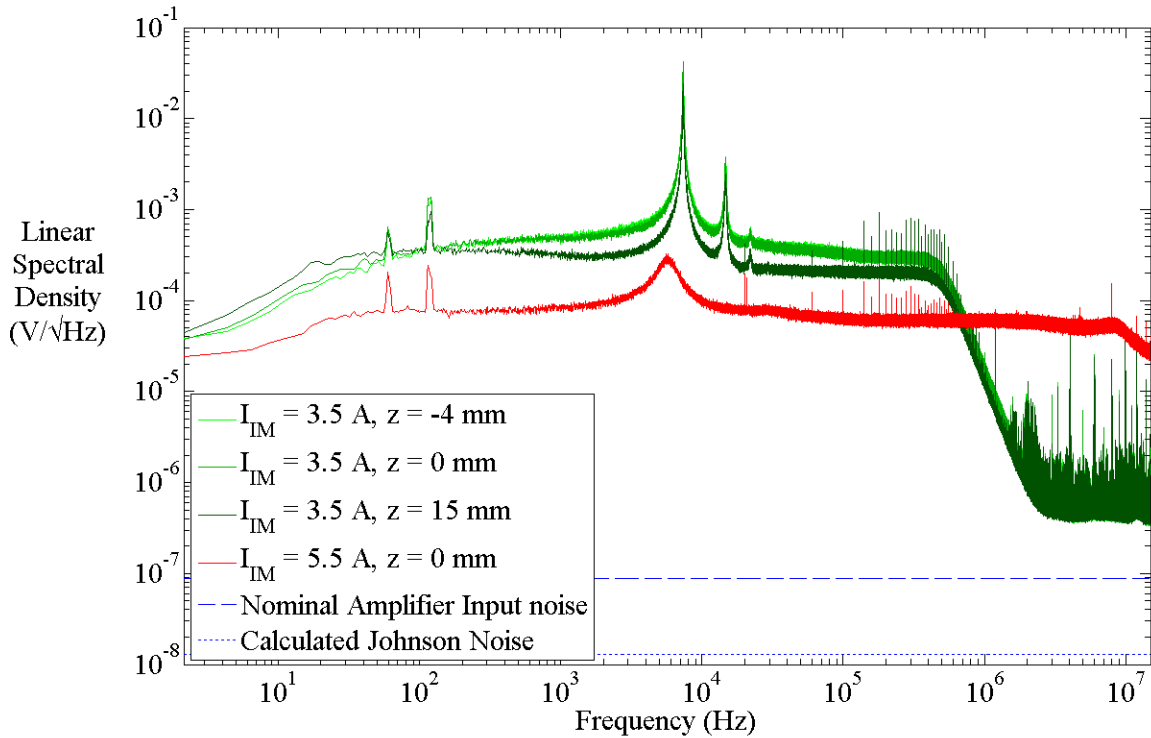


Figure 5.11. Measured noise spectral density of the PMT signal for the H6 operating at 150-V discharge voltage, 10-mg/s anode flow rate. The nonwhite spectra suggest background collected light is the dominant noise source. The measured noise spectra are all orders of magnitude above the nominal amplifier input noise or calculated resistor Johnson noise (shown in blue).

Figure 5.11 shows the linear noise spectral density measured from the H6 operating with a discharge voltage of 150 V and an anode flow rate of 10 mg/s. Separate spectra are plotted in shades of green for three different axial positions at the nominal magnetic field: 4 mm upstream of the exit plane, at the exit plane, and 15 mm downstream. Another line in red shows the spectrum at the exit plane for the high-field condition with an inner magnet current of $I_{IM} = 5.5$ A. The four measured noise spectral densities are plotted together with the calculated Johnson noise for a 1 k Ω resistor at room temperature (300 K) and the nominal input noise of the Krohn-Hite 3945 filter/amplifier, both in blue.

There are three sets of prominent peaks in the noise spectral densities. The most prominent features of the spectra are the low frequency peaks near 10 kHz and harmonics that correspond to breathing oscillations. The peaks at 60 and 120 Hz are likely not physical and come from interference pick-up in the electrical lines downstream of the PMT. There is also a series of peaks beginning near the 100 kHz frequency range and extending up to nearly 1 MHz.

The latter series of peaks notably occurs only at the point 15 mm downstream at nominal field and at the exit plane at high field; they do not occur at the exit plane and 4 mm upstream for the nominal field setting, indicating that they may be related to some process occurring primarily in the plume and not in the discharge channel.

With a generally non-white spectrum and such strong peaks clearly from low frequency plasma oscillations (see the similarity with the discharge current noise spectral density in Figure 7.11), it is clear that noise is dominated by background emission from the plasma, especially at low frequencies.

The noise floor begins rolling-off at about 500 kHz for all of the nominal-field spectra but not for the high-field spectrum. There is a corner frequency near 2 MHz, where the noise density reaches a noisy minimum. The roll-off could be a pink noise region, but it is possible that it is artificial from some unknown low-pass filtering effect. The high-field spectrum was measured the same way and does not feature the roll-off, so a low-pass filtering effect seems unlikely. The only known low-pass filtering effect that might have affected the nominal field settings but not the high-field setting would be human error in setting the Krohn-Hite filter to a cutoff frequency lower than the Nyquist frequency used for the noise spectrum measurements. Further tests are necessary to rule out or verify that possibility.

In any case, the overall trend is of reduced noise spectral density at higher frequencies. That fact, along with the better time resolution available at higher frequencies, makes it advantageous to work at the highest modulation frequency possible, where the noise spectral density is as low as possible. Setting the modulation frequency as high as possible would minimize the power passed by the band-pass filtering and phase-sensitive detection, allowing them to reject much of the noise spectrum and significantly raise SNR before averaging over many cycles of the oscillation.

5.5.1.1 Setting Modulation Frequency to Avoid Local Maxima

The peaks found in the hundreds of kilohertz to megahertz range are significant to be aware of because they are in the range of the modulation frequency. In fact, there happens to be a peak at 499,997.8 Hz, only 2.2 Hz away from 500 kHz, one of the standard modulation frequencies tested in the following sections. Under "normal" circumstances, when the noise spectral density around the modulation frequency is approximately constant, the general shape of the SNR versus f_{mod} plot does not strongly depend on τ . The SNR of each point tends to be

inversely proportional to the transmission window width due to passing more noise with a larger window. Though Figure 5.14 only shows the result for a single time constant, it was actually discovered during the analysis that as the time constant in PSD is reduced (widening the transmission window), the SNR at 500 kHz suddenly drops relative to the other points. The reason for the sudden drop in SNR is apparently that a relatively large increment in noise is passed when the transmission window widens just enough to include the peak.

This scenario illustrates that the best practice for an experiment would be to measure the noise spectral density before the experiment and to specifically choose the modulation frequencies to avoid peaks in the noise spectrum rather than to simply choose "round" numbers such as exactly 500 kHz. Tailoring the modulation frequency to the actual observed noise spectral density may help to improve the SNR, or at least make the shape of the SNR versus f_{mod} profile less dependent on τ . This particular example would not matter for the TFLIF measurement since the series of peaks are densely spaced only 20 kHz apart, and the transmission window for the 2- μ s time constant is about 80 kHz, thus it would be impossible for the TFLIF measurement to avoid peaks. In that case, a time-resolved measurement at the 500-kHz frequency would likely have a lower SNR than a time-averaged SNR test would lead us to believe. There are, however, some small features slightly about the noise in the 1 to 10 MHz range. It would be best practice to find these in the noise spectral density and set the actual modulation frequency to avoid them.

5.5.2 The Raw Signal-to-Noise Ratio

The signal-to-noise ratio (SNR) is commonly defined by other authors as the ratio of the time-averaged signal at the peak of the distribution to the RMS noise (e.g. [45], [62]). This convention is used throughout much of this dissertation except in Subsections 5.5.3-5.5.7, where the average of noise is used instead of RMS to facilitate estimating the uncertainty of SNR measurements. The more common definition with RMS noise has a number of advantages: (1) it is easy to use when the uncertainty of the SNR is not critical, (2) it is more useful to compare with others, and (3) it is more powerful since it can be meaningfully applied to noise with a zero mean.

The raw SNR of the LIF signal can be estimated as follows for the example of the low-field setting of the H6 at the exit plane (see Chapter 7). The time-averaged signal at the peak of the VDF before normalization is 0.0030 V, while the RMS of the digitized signal just following

the band-pass filter is 0.0832 V. Hence for this example, the SNR after filtering and digitization is $SNR_{BP} = 0.0360$, or in other words the noise is about 27.8 times the magnitude of the signal. This requires averaging over tens of thousands of chunks to raise the SNR significantly above unity and, with a chunk size of about 1 ms, this is ultimately what requires the data acquisition to be about 60 s. Note that SNR_{BP} is not actually the raw SNR since it benefits from the significant SNR improvement of the band-pass filter. The SNR improvement of the band-pass filter is about $I_{BP} = 50$, and therefore the raw SNR is approximately $SNR_{raw} = 7.2 \times 10^{-4}$, meaning the noise amplitude is about 1400 times the signal amplitude in the raw photomultiplier signal. The estimate for this example is typical for the raw SNR of LIF data.

5.5.3 Measuring the Signal-to-Noise Ratio

Throughout this dissertation, but especially in this chapter, it is often of interest to measure the signal-to-noise ratio as a function of various parameters to determine optimal parameter settings. Time-resolved measurements are highly expensive in terms of experimental and computational time, so these SNR tests are carried out on time-averaged LIF signal using the TFLIF hardware under the same conditions for a time-resolved measurement, but with a much shorter investment of time (both acquisition and computation time) and storage space.

The SNR is measured by phase-sensitive detection performed in software set to a long time constant of about 100 ms with the laser wavelength set to the peak of the LIF profile. The SNR is estimated by n measurements with the laser on followed by n measurements with the laser off and is defined in this case by:

$$SNR = \frac{\langle A_{laser\ on} \rangle - \langle A_{laser\ off} \rangle}{\langle A_{laser\ off} \rangle} = \frac{\langle A_{laser\ on} \rangle}{\langle A_{laser\ off} \rangle} - 1, \quad (5.11)$$

where $A_{laser\ on}$ is the PSD signal amplitude when the laser is on, $A_{laser\ off}$ is the PSD signal amplitude when the laser is off, and brackets denote an average over the n measurements ($n = 80$ was used in most cases). The reference signal is locked to the effective signal phase offset using that the same phase locking code developed for TFLIF so that the PSD signals are the magnitude of the signal or noise. This is equivalent to the magnitude from the vector computer technique using two arbitrary, orthogonal reference signals.

Similar to the treatment of noise by most authors, this definition implicitly assumes that the noise does not depend on the laser frequency so that it can be interpreted as the signal-to-

noise ratio at the peak of the distribution profile. Of course, the SNR is obviously lower at points far from the peak where the signal is much lower but the noise is approximately the same. Smith provided evidence that noise does actually change significantly depending on the laser frequency [62], especially for frequencies where the laser interacts with the plasma to generate an LIF signal (see Figures 5.1 and 5.2 in the reference). Even so, the ratio of the LIF signal at the peak of the distribution to the noise at the edges of the distribution (or when the laser is off entirely) is a simple and useful way to quantify the SNR with a single number, even if that number does not strictly correspond to the actual SNR at the peak or any other point in the distribution.

Equation (5.11) is a somewhat uncommon definition; many authors define SNR in terms of a signal divided by an RMS noise. The reason to use the uncommon definition here is that we measure how the SNR changes as a function of varying parameters to help optimize the parameters, and the uncertainty in the SNR calculation should be known to help evaluate whether a given change in SNR is significant or possibly merely result of random variation. Most authors using the common definition neglect the uncertainty of the SNR measurement. With SNR defined in terms of the average signal divided by average noise, it is simple to calculate the standard error in the two factors and therefore the standard error in the SNR by propagation of uncertainty. If SNR were defined using the RMS noise, then the n noise measurements could be used to calculate an estimator of the RMS noise, but we would need a number of sets of n measurements to estimate the standard error of the RMS estimator. This definition then simplifies both the experiment and the calculation of the uncertainty in SNR.

The subtraction of 1 is also uncommon but is only a matter of taste. It ensures that the $\text{SNR} = 0$ when the output PSD signal is the same with the laser on as it is with the laser off, as it should be. If there is no detected signal but only noise, then surely the signal-to-noise ratio should be 0 and not 1. Of course, this shift does not affect the uncertainty statistics or the interpretation of the results.

A critical question to answer is whether this is an appropriate definition to quantify SNR. RMS is often used in a variety of applications related and unrelated to LIF perhaps because random noise tends to have zero mean, and in that case a definition of SNR using mean noise would clearly not be appropriate. In the case of the magnitude of signal from PSD, however, the noise and signal are always positive and the average tends to scale with the RMS, and therefore the average noise can be, in a sense, a proxy for the RMS. The SNR values calculated with this

method are not in general equal to the more common definition since RMS is more sensitive to outliers. In practice, there are rarely outliers far from the mean noise in LIF datasets, so they scale similarly.

5.5.4 SNR vs. Laser Modulation Frequency with the Hollow Cathode

For plasma conditions typically associated with a Hall thruster or hollow cathode, the signal photon rate (i.e. the rate at which LIF photons are captured) can be a limiting factor in time-resolved measurements. In fact, Mazouffre et al. have even opted to use a photon counting technique and they argued that the low signal photon rate requires it [75]. The signal and noise photon rates can vary widely depending on the plasma source and location interrogated, the laser power, size of the interrogation volume, and the solid angle subtended by the collection optics. Therefore, different systems may collect very different photon rates despite the nominally similar application of "LIF in a Hall thruster."

The collected signal photon rate can limit the maximum usable modulation frequency for an analog signal measurement technique. As modulation frequency increases, there are fewer signal photons per modulation period. The signal photon rate eventually falls too low for the digitizer to resolve a smooth analog signal in the PMT signal, even with the band-pass filter smoothing photoelectron pulses into an analog signal. The SNR of an analog analysis technique will suffer and possibly fail to detect signal at all. This is a fundamental limit and is the mechanism suspected to cause the modulation frequency limit that is currently observed.

We tested the SNR versus modulation frequency in the experimental setup with the hollow cathode in Chapter 6 in two cases: (1) with unfiltered collected light; and (2) with a 0.3 OD neutral density filter to block about 50% of collected light. Removing the filter and doubling the collected light has different effects on the signal and noise. The signal amplitude doubles since the signal light adds coherently, but the random, incoherent noise amplitude will only increase by $\sqrt{2}$. Assuming background light dominates all sources of noise, we expect the ratio of unfiltered SNR to filtered SNR to be $SNR_{unfiltered}/SNR_{filtered} = \sqrt{2}$ unless this process also changes the signal or makes other noise sources significant.

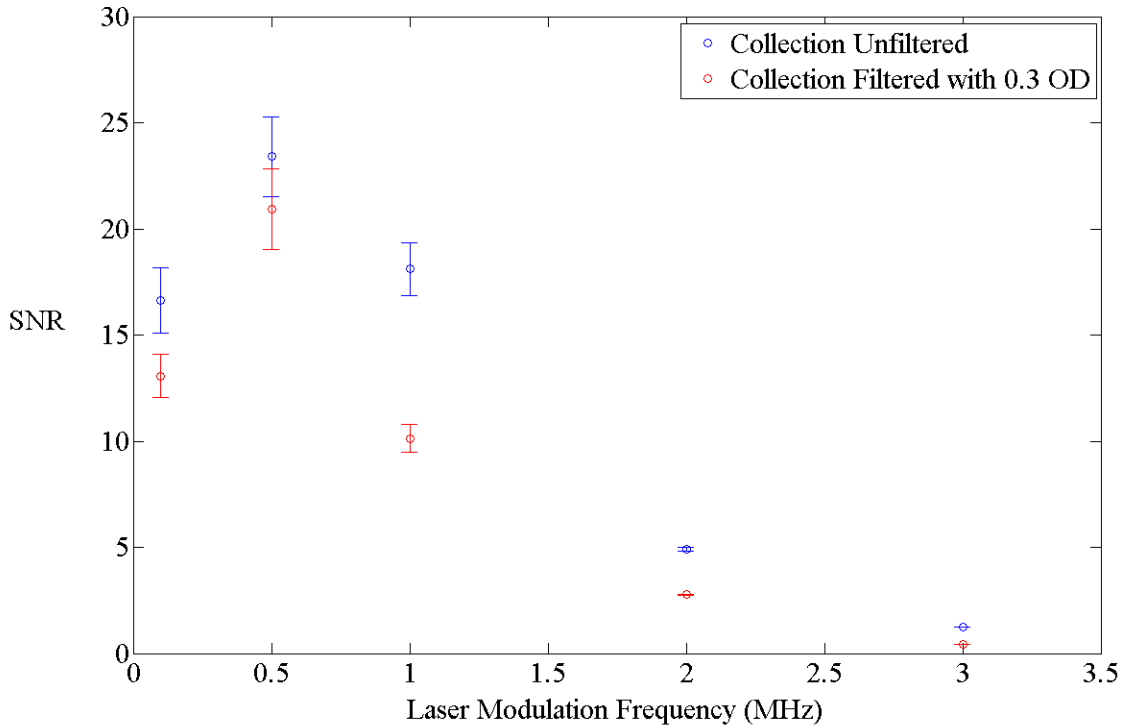


Figure 5.12. Signal-to-noise ratio of time-averaged LIF signal reaches a maximum near a laser modulation frequency of 500 kHz and falls toward zero thereafter. Filtering about 50% of collected light dramatically reduces SNR by more than 50% at high modulation frequencies, but only by 20% at most at the low modulation frequencies.

The results of this test for the hollow cathode test bed are shown in Figure 5.12. At modulation frequencies up to 100 and 500 kHz, the filter only moderately affects SNR, with an SNR improvement ratio of about $SNR_{unfiltered}/SNR_{filtered} = 1.1 - 1.28$. The filter actually affects the SNR less than expected at these frequencies, but the relatively high uncertainty may explain it. The SNR reaches a maximum near $f_{mod} = 500$ kHz before rolling off relatively sharply at modulation frequencies on the order of megahertz. For those megahertz-range frequencies, the filter affects SNR more strongly with a ratio of about $SNR_{unfiltered}/SNR_{filtered} = 1.8 - 2.9$.

5.5.4.1 Interpretation of Results

The increased ratio at the higher frequencies indicates that another effect influences the SNR beyond the coherent addition of signal and incoherent addition of noise. There are at least two possible culprits; unfortunately, this experiment does not fully isolate possible causes. One possibility, as mentioned in the first two paragraphs of this subsection, is that signal photon rate

per modulation period may be falling too low to produce a clean analog signal at higher modulation frequencies.

Another possible explanation comes from the fact that the band-pass filter bandwidth was not held constant. During this test, the cutoff frequencies of the pass band were set to $f_{mod} \pm 0.1f_{mod}$ for the frequencies up to 2 MHz. The pass band was 2 – 3.3 MHz for $f_{mod} = 3$ MHz due to a limitation of the Krohn-Hite filter. In any case, the pass band increases monotonically with the modulation frequency. Then the band-pass filter passes more noise, and hence the SNR of the raw voltage signal at the digitizer is dramatically worse at higher modulation frequencies. From the discussions of Subsections 4.3.1 and 4.3.2 and the empirical result of Subsection 5.5.7, ideally, a lower I_{BP} should be compensated by an increased I_{PSD} since the phase-sensitive detector acts as a narrower band-pass filter that removes noise components left over by the band-pass filter. That analysis, however, does not take into account the quantization noise that could result if the signal level is much lower than the noise amplitude and not well resolved by the digitizer (see Subsubsection 5.1.1.2). It is possible that the increased noise at higher modulation frequencies put the system into a regime where quantization noise degrades the SNR. This issue could be largely removed by keeping the filter pass band constant, though the noise amplitude would still change somewhat with the noise spectral density in the region near the modulation frequency.

Unfortunately, the small number of points available hampers the interpretation. Testing even lower modulation frequencies and using a denser selection of points would help to make trends clearer and also enable detecting a change in the location of the maximum. It is not clear, but a change in the frequency of the maximum is already hinted by the points available.

5.5.4.2 Implication for TFLIF experiments

Figure 5.12 demonstrates that we were unable to acquire data with the cathode source above a modulation frequency of about 2 MHz, and even modulation at 1 MHz comes at the cost of reduced SNR. A 1-MHz modulation frequency was the standard for the campaigns in the later chapters of this dissertation because it is a good tradeoff between SNR and time resolution.

The hope at this stage was that if the signal photon rate per modulation period were the cause of the reduced SNR at higher modulation frequencies, then the higher ion density and electron energy in a Hall thruster may lead to a higher collected signal photon rate and therefore a higher modulation speed limit. If the collected photon rate is indeed the problem, then it is a

limit fundamentally linked to the properties of the plasma source, though it could be improved with a higher laser power and collection optics subtending a larger solid angle.

Note that the data presented in Figure 5.12 come from a more recent analysis of the same raw data used for FIG. 7 in a journal paper [89]. The two figures show the same qualitative behavior, but are slightly different due to minor changes in the PSD algorithm used and different parameters for time constant and the number of samples used in the average. Increasing the number of samples tends to reduce the standard error, but it also reduces the SNR by reducing the time constant of each sample when the samples are taken from a dataset of fixed acquisition time. The actual value of SNR is not critical and will change with the time constant and other parameters, but the trend of SNR as a function of modulation frequency is important.

5.5.5 The Effect of AOM lenses and Sampling Frequency

Following the above test, it was suspected that the output laser modulation waveform from the AOM might not be following the input waveform as well as could be expected given the nominal rise time for the 200-mm-focal-length lenses. A new photodiode with a bandwidth of 180 MHz (Thorlabs PDA10A) was purchased to ensure a more accurate measurement of the modulation waveform compared with the old photodiode with a bandwidth of 10 MHz (Thorlabs PDA36A). Also, the 200-mm focusing and collimating lenses of the AOM were replaced with 120-mm lenses, giving a nominal rise time of 19 ns, down from 32 ns for the 200-mm lenses. The SNR versus modulation frequency was then tested again.

In addition, since the number of samples per modulation period falls as modulation frequency increases at a fixed sampling frequency, it is also expected that SNR as a function of modulation frequency will eventually be limited by the sampling frequency when the digitizer no longer sufficiently resolves the laser modulation waveform for phase-sensitive detection. The test with 120-mm lenses was performed at sampling frequencies of both 20 MHz and 30 MHz to detect any significant changes in the SNR that depend on sampling.

The result is shown in Figure 5.13. The maximum now occurs at 1MHz instead of 500 kHz in the coarsely sampled frequency space. This result indicates that the laser modulation waveform from the AOM with the 200-mm lenses may indeed have been a limiting factor for the modulation frequency but not necessarily the only one. It was unclear at this point if the AOM response with 120-mm lenses could still be a limiting factor.

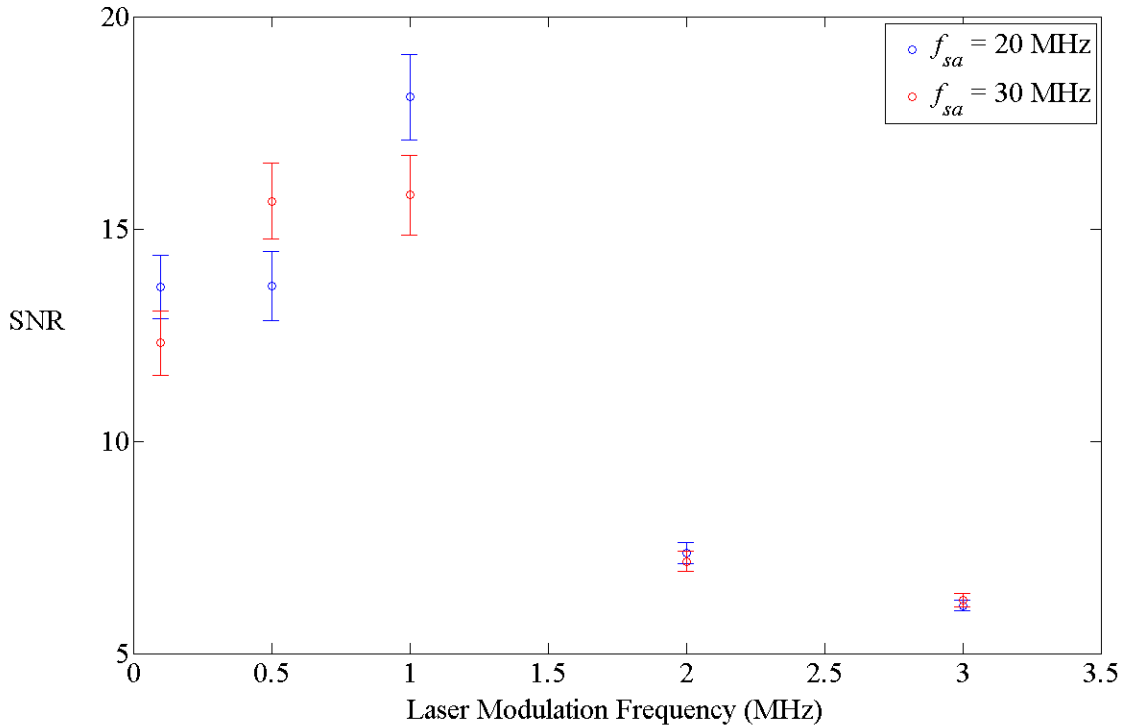


Figure 5.13. AOM lenses with a 120-mm focal length allow the SNR to reach a maximum near a modulation frequency of 1 MHz, significantly higher than with the 200-mm lenses of Figure 5.12. The SNR is insensitive to the sampling frequency at these conditions.

In addition, the results show that the SNR appears to be insensitive to the sample frequency, at least within the uncertainty of the measurements. The 30-MHz sample frequency had a higher SNR at some points but not at others with no discernible pattern. This suggests that the 20 and 30 MHz sample frequencies both sufficiently resolve the modulation waveform for phase-sensitive detection at all of the frequencies tested. It may be possible to achieve similar results with an even lower sample frequency, saving on data storage and computation time.

5.5.6 SNR vs. Laser Modulation Frequency with the H6

Finally, a test of SNR versus modulation frequency was performed with the H6 as part of the initial campaign described in Chapter 7. The collection optics and interrogation volume were the same as well as other parameters such as sampling frequency (30 MHz) and filtering. The results in Figure 5.14 show that the SNR has a maximum near a modulation frequency of 2 MHz. Barring some unknown parameter that could explain the change, the change of the location of the

maximum suggests that the steep drop in SNR is determined by the plasma properties as opposed to the hardware or software of the measurement.

Unfortunately, the result is potentially consistent with either hypothesis for the origin of the decline in SNR as a function of modulation frequency. The location of the maximum would be expected to move if the collected signal photon rate changed and the origin of the decline is a low signal photon rate per modulation period. If quantization noise is the issue, the location of the maximum could also change since the background emission is strongly dependent on electron temperature. The source could be isolated by performing the SNR versus modulation frequency test with a constant filter bandwidth.

If the main limiting factor is the collected signal photon rate, then the implication is that the modulation frequency cannot be greatly improved by upgrading hardware, and it is fundamentally limited by the plasma source. If so, then the H6 is limited to a modulation frequency of a few megahertz at the operating condition of 150-V discharge voltage and 10-mg/s flow rate, and therefore the time resolution is limited to the order of hundreds of kilohertz. This is a sufficient bandwidth to interrogate breathing, spoke, and possibly even high frequency cathode oscillations near 100 kHz, but higher frequency oscillations will likely remain out of reach of the TFLIF system. There may be some flexibility to collect light from a larger solid angle, reduce losses, raise laser power etc., but such efforts are not likely to raise the collected signal photon rate by orders of magnitude.

If the LIF signal photon rate is the limit, then the maximum in the SNR as a function of modulation frequency will vary in space with the ion density and electron temperature (since they largely determine the density of the metastable state probed). For that reason, the SNR test was performed with the H6 at two spatial points: at 4 mm upstream of the exit plane ($z = -4$ mm) and at 15 mm downstream of the exit plane ($z = 15$ mm). The SNR is generally lower at $z = 15$ mm as expected. No major shift in the location of the maximum was detected, but that may be due to the coarse resolution of frequency space. The plasma conditions are very different at the two locations [69], so it is unexpected that the maximum apparently occurs near same frequency. If anything, the relatively high point at 3 MHz suggests the maximum and roll off occur at a higher frequency at the downstream point. This counterintuitive result undercuts the evidence in favor of the photon rate limit and is potentially consistent with a limit by either quantization

noise or the AOM itself. Clearly, more work is necessary to uncover the causes of the modulation frequency limit.

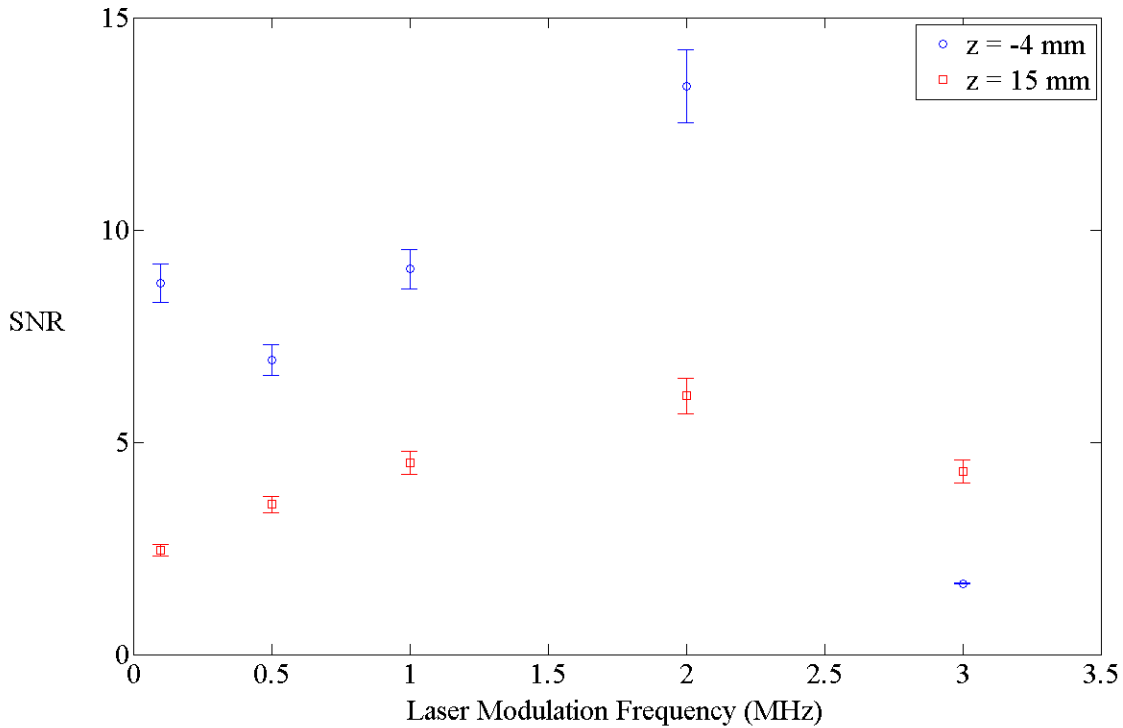


Figure 5.14. The SNR with the H6 reaches a maximum at the laser modulation frequency of 2MHz. The SNR is much lower at the point 15 mm downstream of the exit plane than at 4 mm upstream, but the maximum occurs at the same frequency for both.

5.5.7 Electronic Band-Pass Filter

The effect of filter bandwidth was tested empirically early in the development of the TFLIF system. Figure 5.15 shows the results of a test with the hollow cathode setup of Chapter 6 and a modulation frequency of $f_{mod} = 100$ kHz. As predicted by the discussion of Subsections 4.3.1 and 4.3.2, the SNR does not depend strongly on the filter bandwidth, if at all within the noise of the SNR measurements. The filter bandwidth given on the x-axis is the conventional definition of bandwidth for a band-pass filter: the difference between the upper and lower cutoff frequencies. All of the pass bands tested were centered on the modulation frequency. For the bandwidth of 100 kHz, for example, the cutoff frequencies are 50 kHz and 150 kHz.

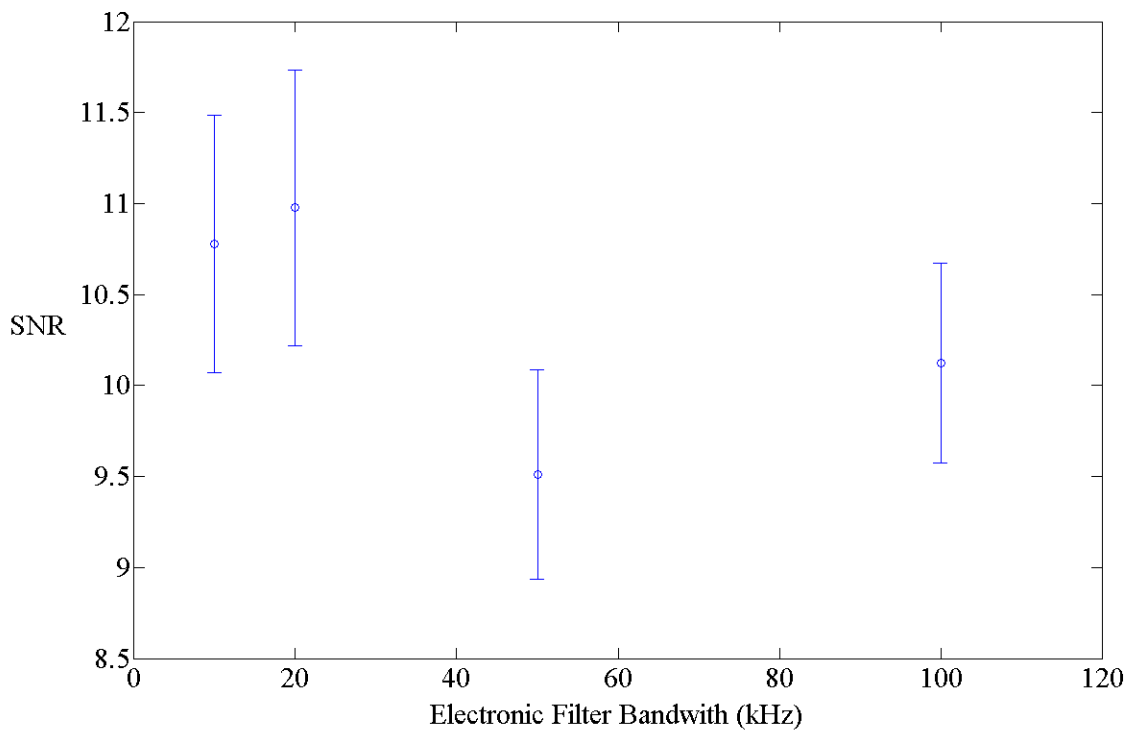


Figure 5.15. SNR as a function of electronic filter bandwidth with $f_{mod} = 100$ kHz. SNR does not depend strongly on the electronic filter bandwidth.

Chapter 6

Validation Experiments for TFLIF with a Hollow Cathode

And we daily in our experiments electrise bodies plus or minus, as we think proper. [These terms we may use till your Philosophers give us better.] To electrise plus or minus, no more needs to be known than this, that the parts of the Tube or Sphere, that are rubb'd, do, in the Instant of Friction, attract the Electrical Fire...

– Benjamin Franklin (Letter to Peter Collinson, 25 May 1747)

Following the theoretical development of the TFLIF technique in Chapter 4 and the progression toward a practical implementation of the technique in Chapter 5, the goal of this chapter is to validate that the system is capable of accurate measurements. This chapter describes the validation campaign using a hollow cathode with controlled discharge current oscillations. We have designed a series of tests that, when taken together, provide strong evidence for the general validity of the technique and for the accuracy of the measurements from the system implemented at PEPL.

Section 6.1 introduces the vacuum facility and the plasma source, and describes the details of the experimental setup with which validation tests were performed. We briefly note the qualitative significance that the TFLIF system captures a signal if and only if using the correct parameters in Section 6.2 before moving on to the individual validation experiments. Results from an experiment using a periodic signal are presented in Section 6.3 as a consistency check that TFLIF returns the same results as TALIF. The triggered average fails for non-periodic signals, in which case the TFLIF and TALIF signals cannot be expected to converge. Section 6.4 presents an experiment with random oscillations to check that the transfer function reproduces the failed TALIF signal when using the triggered average of discharge current as the input signal. The transfer function's ability to reproduce the triggered average signal was shown in Subsection 4.3.5, where comparing the actual triggered average to the transfer function's reproduction was proposed as a general purpose benchmark to validate the transfer function for all signals. Finally, Section 6.5 summarizes the results and validation arguments.

6.1 Facility and Plasma Source

The experiments presented in this chapter were performed in the "Junior" or "Foster" chamber at PEPL (see Figure 6.1), which is a 1-m-diameter by 3-m-long chamber connected to the much larger (6-m by 9-m) Large Vacuum Test Facility via a gate valve. Junior has a Leybold MAG W 2010 C turbopump as its primary pump, providing a base pressure of 2×10^{-6} Torr and a background pressure during cathode operation of 4×10^{-5} Torr.

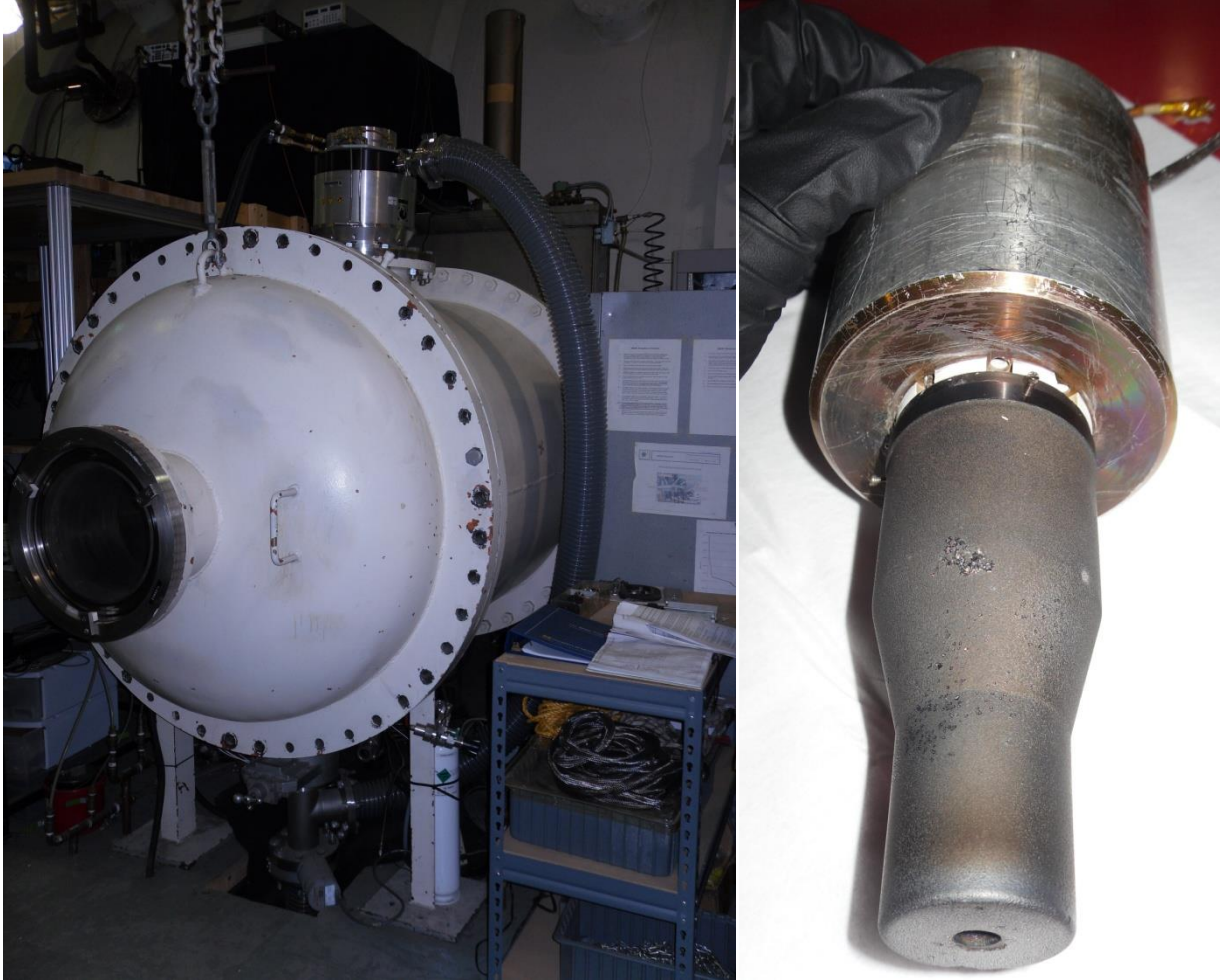


Figure 6.1. Photographs of the "Junior" Chamber at PEPL (left) and the cathode used in the validation campaign (right).

The plasma source is an orificed hollow cathode [21], [22] that was originally designed as the ionization stage cathode of the NASA-173GT [99], a two-stage hybrid Hall/ion thruster. It was custom made by the Busek Corporation and has a nominal discharge current up to 60 A. Though the cathode has a nominal 10-sccm flow rate, the LIF signal was not usually measurable at 10 sccm. The following experiments operated between 3.5 and 7 sccm to maximize the LIF

signal. Collisional quenching of the metastable state in higher background pressure may have caused the loss of LIF signal at higher flow rates, but that has not been substantiated.

To establish a highly controllable discharge current oscillation, the experiment used a pair of anodes powered by bipolar power amplifiers (a Kepco 100-2M and a Kepco 100-4M) in current control mode. The anodes are 7.6-cm-diameter rings made of 0.635-mm-thick stainless steel. Anode 1 is 5 cm long, positioned 10 cm from the cathode keeper plate, and supported by the 2-A supply. Anode 2 is 10 cm long, positioned 18 cm from the cathode, and supported by the 4-A supply. The "discharge current" reported here is the total discharge current, or the sum of currents in both anodes. Two anodes were used with separate power amplifiers (but a common driving signal) because a single high-speed amplifier could not supply sufficient current for a stable discharge and strong LIF signal. The discharge current was found to be extremely stable in the range of approximately 4-6 A. With the cathode operating well below its nominal power, the heater and keeper were left on during operation to help stabilize the discharge.

A diagram of the experimental setup is shown in Figure 6.2. The configuration of the optics is the same for all experiments in this dissertation and was detailed in Section 4.2. The interrogation volume is a fixed location at the edge of the keeper plate concentric with the hollow cathode orifice, chosen to maximize SNR.

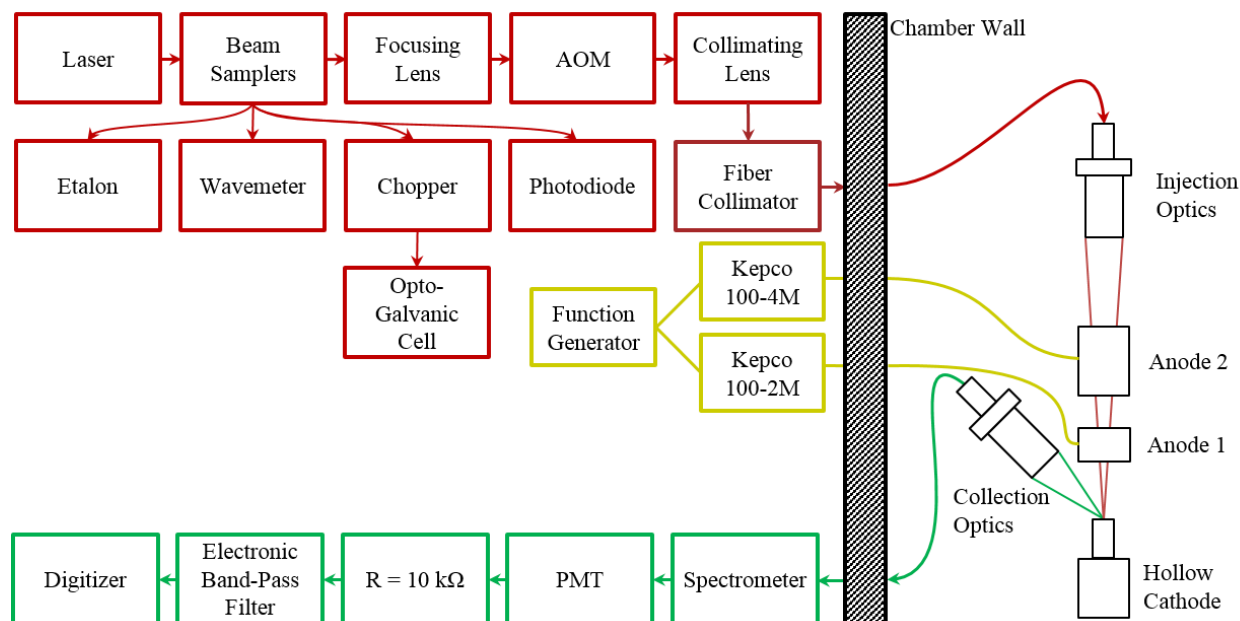


Figure 6.2. Diagram of the optics and plasma source used in the validation campaign. Laser light follows the red path, collected fluorescence light follows the green path, and the connections of the power amplifiers to the anodes and function generator are shown in yellow.

One reason to use a hollow cathode instead of a Hall thruster for validation is to control the discharge current more easily than is possible with a Hall thruster. This control allows two validation experiments that would be impossible with a Hall thruster. The first experiment uses a periodic oscillation. If the discharge current oscillates periodically, then the oscillation of plasma parameters, such as the ion velocity distribution and therefore LIF signal, should be periodic as well, turbulent and stochastic effects notwithstanding. That allows the easy use of triggered averaging with the discharge current as a trigger. Comparison of the TFLIF result to the triggered average for a periodic signal is the subject of the first experiment in Section 6.3. The second experiment, the subject of Section 6.4, uses a random oscillation in the sense that each new cycle of the sinusoidal oscillation has a random period. Demonstrating that the TFLIF technique faithfully recovers the signal even for a nonperiodic and random signal is key to credibly applying the technique to Hall thruster oscillations. Since the triggered average no longer applies, this experiment gives an opportunity to test the benchmark proposed in Subsection 4.3.5.

6.2 Preliminary Validation Considerations

Before delving into the details of the individual tests, we note one important feature of the measurements of the TFLIF system. Analysis yields a significant TFLIF signal only if it is applied at the correct reference frequency to a data set captured when the laser interacts with the plasma. Also, if analysis is applied at an incorrect reference frequency or if the laser intensity is zero, then the "signal" returned by this analysis technique is at least an order of magnitude lower than a significant TFLIF signal and its waveform resembles white noise. In other words, the technique measures a signal if and only if the conditions are correct for a signal to be measured.

This is clearly a necessary feature of any accurate measurement, but it is important to note this criterion here because we apply a complicated and unusual analysis technique that could potentially be susceptible to some form of artifact. For example, the discharge current signal is used in the transfer function averaging and characteristic output synthesizing processes, and TFLIF signals tend to have a similar Fourier spectrum as the discharge current, hence it is plausible that there might be some artifact that could cause the output LIF signal to have similar frequency components as the discharge current even if there is no actual LIF signal. The fact that the system measures a signal only under the proper conditions implies that there is no such artifact in analysis that causes a false LIF signal in the output, a necessary step before any validation experiment.

6.3 Validating the TFLIF system with Sinusoidal Discharge Current Oscillations

This section presents an experiment with 10-kHz sinusoidal discharge current oscillations that was performed on 11 March, 2013. LIF data are captured at 16 wavelengths so that the VDF profiles can be compared to diagnose possible distortion of the VDF profile by the transfer function averaging technique. A frequency of 10 kHz was used to approximate the frequency of Hall thruster breathing mode oscillations. Subsection 6.3.1 discusses the motivation and goals for the experiment and Subsection 6.3.2 presents the pertinent results, primarily a comparison between the TFLIF signal and the triggered average LIF signal.

The operating conditions were selected as follows to maximize the SNR of the LIF signal. The gas flow rate was 3.5 sccm. The discharge current oscillated between about 4.25 A and 6.1 A (see Figure 6.3), while the discharge voltage was about 33 V on anode 1 and 43 V on anode 2. The same 10-kHz sinusoidal control signal was used for both current amplifiers, provided by an Agilent 33220A function generator.

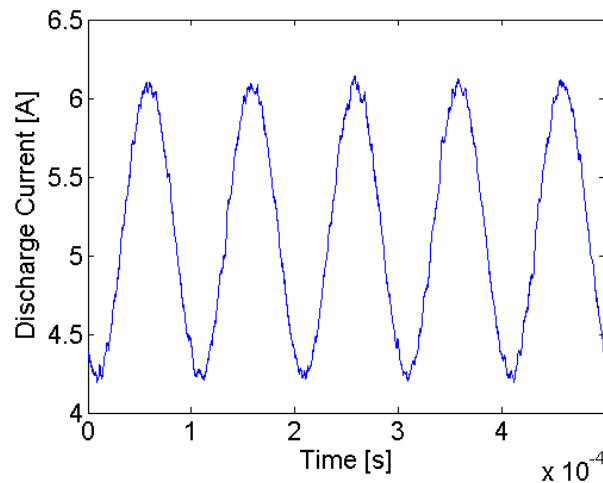


Figure 6.3. The discharge current oscillated between about 4.25 A and 6.1 A with a sinusoidal oscillation fixed at 10 kHz.

6.3.1 Motivation for an Experiment with Sinusoidal Discharge Oscillations

The basic strategy in validation is to compare TFLIF signal with other measurements that it should theoretically agree with under certain conditions. The most obvious possibility is to compare the time-averaged TFLIF signal with the traditional time-averaged LIF signal measured with a lock-in amplifier. This comparison is not sufficient because the time-resolved features of the VDF profile could, in principle, be distorted even if the average values are accurate. The transfer function averaging technique in particular is uncommon and probably is the part of

analysis that is most open to doubt. Therefore, we validate time-resolved features in addition to the average signal comparison

A sinusoidal signal is convenient because it is periodic and therefore the triggered averaging technique will converge to the actual signal (see Subsection 4.3.4). Meanwhile, the transfer function average will converge to the actual signal if the output is related to the input by a time-invariant, linear system (Subsection 4.3.3). We can verify that the transfer function averaging technique is not introducing systematic error to time-resolved features by showing that the two signals agree for a periodic oscillation.

Both the TFLIF and triggered averaging are performed on the same data set by software in post-processing. Therefore, the laser modulation and analog PMT signal filtering are the same. In order to demodulate the signal, both analyses use phase-sensitive detection with the same time constant of 2 μ s. The comparison isolates the effect of the averaging technique because the two analyses differ only in that one used the transfer function averaging technique following PSD, and the other applied triggered averaging (see Figure 4.2). For triggered averaging, each time-series is triggered at the phase corresponding to the maximum in the discharge current after smoothing to reduce effects of noise. The smoothing used in this case is a zero-phase digital first-order Butterworth low-pass filter with a cutoff frequency of 10 kHz.

6.3.2 Results

We begin by comparing a time average of the ion VDF profiles from both time-resolved techniques with the time-averaged LIF profile measured with a commercial lock-in amplifier; i.e. conventional time-averaged LIF data. An example of this comparison for the cathode experiment is shown in Figure 6.4. The profiles have been normalized to unity peak magnitude, with the (noisier) lock-in amplifier signal normalized by the maximum of its profile smoothed by a 31-point moving average. All three profiles match within the error of the lock-in amplifier measurement, which is roughly apparent because there are many closely spaced points. This confirms that the average LIF profiles measured with the time-resolved techniques agree with the conventional measurement of a lock-in amplifier.

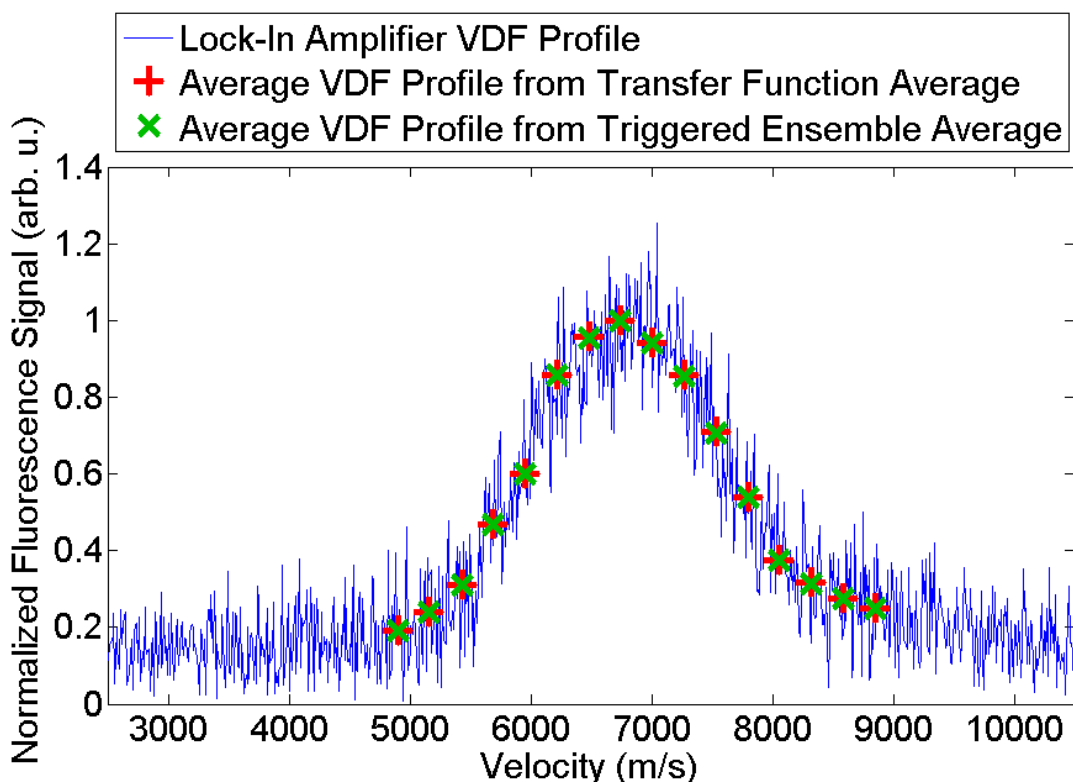


Figure 6.4. A comparison of the average of time-resolved VDF profiles from transfer function averaging (red "+") and triggered averaging (green "x") shown against the average VDF profile measured by the lock-in amplifier LIF system (blue line).

The average normalized TRLIF signals are so close at all wavelengths that all of the markers shown in Figure 6.4 overlap. Different marker shapes were chosen to make the bottom markers visible. The mean absolute residual (the average of the absolute value of the difference between the two TRLIF measurements) is 0.0012. A very small value can be expected due to averaging over 60 s of data, far more than a typical lock-in amplifier measurement with a time constant of about 100 ms. It also implies that the transfer function average does not introduce systematic error in the average LIF profile, otherwise the profiles would not converge so well.

Heat maps of the two LIF signals and their residual (the difference between the two) as functions of time and velocity are shown in Figure 6.5. The two signals are close at all times. There is no obvious systematic difference, and the residual appears to be due to only the random noise in each of the measurements. The mean absolute residual is 0.07, or 7% of the peak value in LIF signal.

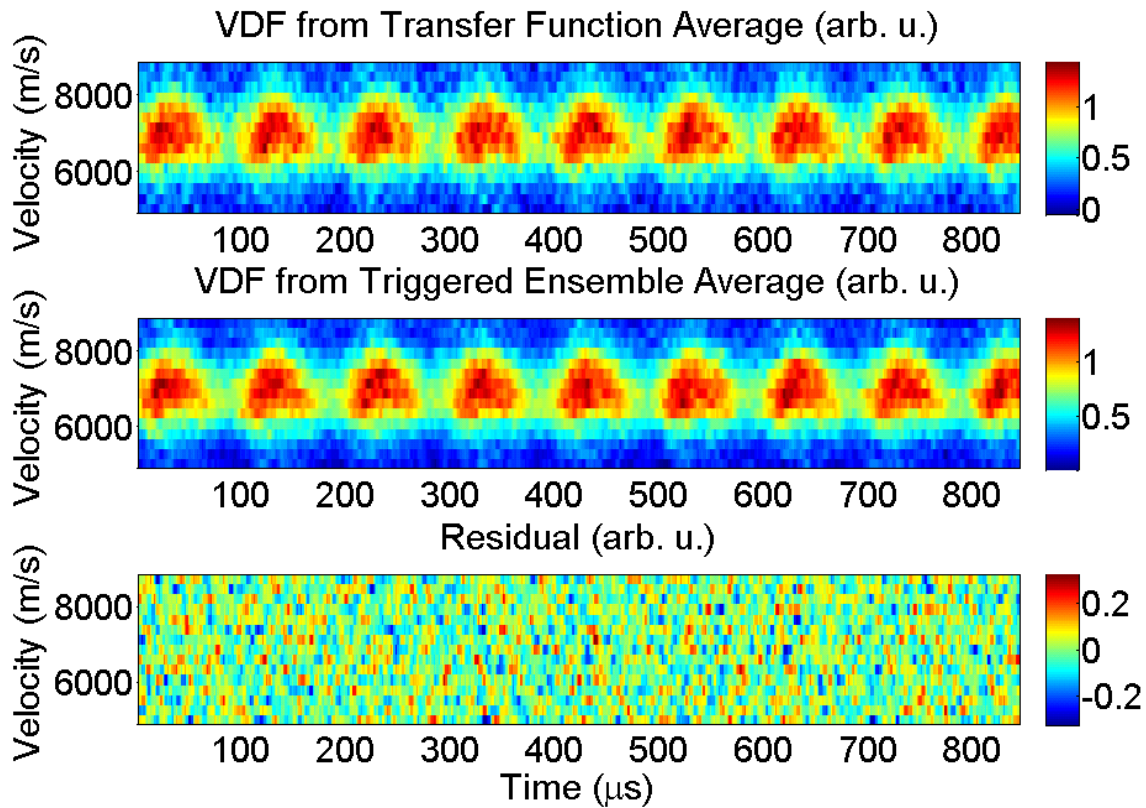


Figure 6.5. Heat maps of the TFLIF (top) and TALIF (middle) signals and their residual (bottom) as functions of velocity and time show no systematic differences between the two signals.

Figure 6.6 is a plot of two "snapshots" of the same ion VDF from Figure 6.5 to more clearly show the shape and behavior of the VDF profile. The VDF profile is somewhat coarse since there are only 16 points in velocity, but it is clear nonetheless that both techniques capture the same general features such as the mean and spread of the profile. There is a slight acceleration of the mean velocity with increasing discharge current.

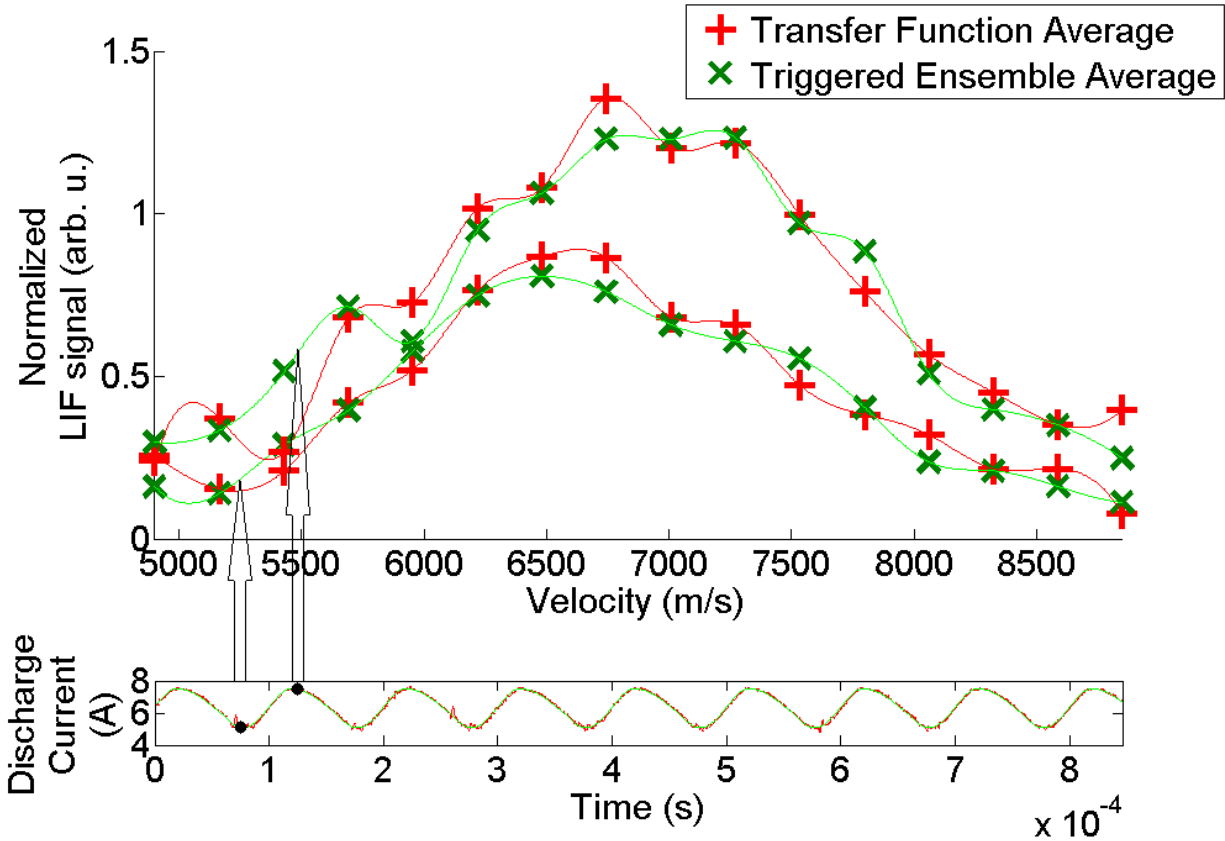


Figure 6.6. "Snapshots" of the VDF profile in time for the transfer function average (red "+") and triggered average (green "x"). The solid lines are spline interpolations meant to be a guide to the eye without a precise physical meaning.

Since the two averaging schemes result in nearly identical TRLIF signals, we conclude that the TFLIF scheme is likely not producing systematic error and is working correctly under these conditions. If there is some systematic error, then either it is in the raw data and does not originate from the averaging scheme, or both averaging schemes somehow introduce the same systematic error, which is unlikely due to their completely independent algorithms.

6.4 Validating the TFLIF system with Random Discharge Current Oscillations

This section focuses on a validation experiment of the TFLIF system with random oscillations on a hollow cathode discharge that was carried out on 21 February 2014. The need for such an experiment to go beyond validation for periodic oscillations is discussed in Subsection 6.4.1. Subsection 6.4.2 discusses how the random oscillations are generated using the "List Mode" feature of a National Instruments function generator with a list of random frequencies. Subsection 6.4.3 specifies the distribution of the random variable used to generate

the frequency list and the simulations performed to determine appropriate parameters for the distribution. Subsection 6.4.4 calculates properties of the expected triggered average waveform as a consistency check for the simulation results. The advantages and limitations of using List Mode compared to other options such as defining an arbitrary waveform are discussed in Subsection 6.4.5. Finally, the results of this experiment in the cathode discharge are presented in Subsection 6.4.6, where measurements of the nonperiodic oscillation are validated in part by applying the benchmark proposed in Section 4.3.5.

The operating conditions were selected to maximize the SNR of the LIF signal. The gas flow rate was 7 sccm. The discharge current oscillated between about 5.5 A and 8.5 A (see Figure 6.7), while the discharge voltage was about 20 V on anode 1 and 26 V on anode 2. The same sinusoidal control signal with a randomly varying period in the range of 7.5 - 12.5 kHz was used for both current amplifiers, provided by an NI PCI-5406 function generator (see Subsections 6.4.2 and 6.4.3 for details).

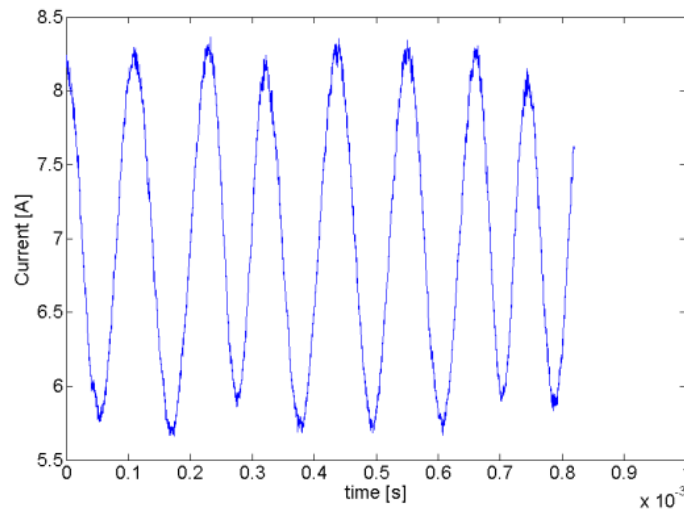


Figure 6.7. The discharge current oscillated between about 5.5 A and 8.5 A with a sinusoidal oscillation with a randomly varying period in the range of 7.5 - 12.5 kHz.

6.4.1 Motivation for an Experiment with Random Discharge Oscillations

The experiment with a sinusoidal oscillation in a hollow cathode discharge provides evidence validating the TFLIF system, but the experiment lacks a key result. We require a demonstration that the system accurately recovers the TRLIF signal even with a non-periodic, non-repeatable oscillation. In other words, we need a demonstration under conditions more similar to the Hall thruster oscillations that are ultimately of interest. After this deficiency was

mentioned by Stéphane Mazouffre at the 2013 IEPC in Washington, DC, an experiment was devised to test the TRLIF system using a controlled but non-periodic oscillation in the discharge current. The goal of this final validation experiment is to demonstrate accurate signal recovery with the transfer function averaging technique while at the same time demonstrating the failure of the triggered average for this random, non-periodic oscillation.

6.4.2 Generating Random Oscillations

The main challenge to the experimental setup was arranging for the non-periodic yet controlled oscillation. The experimental apparatus is exactly the same as the previous hollow cathode experiment except that an NI PCI-5406 function generator drives the anode amplifiers instead of the Agilent 33220A.

One of the advanced functions of this function generator is the so-called "List Mode," which enables the seamless generation of a sequence of waveforms with lists of both frequency and duration. The waveform's frequency begins with the first element in the frequency list for a duration given by the first element in the duration list. It then moves on to generate another frequency for a certain duration given by the second elements in the lists, and so on until the lists are exhausted. The output waveform is always a continuous function because the next step begins at the same phase where the last step ended and because the function generator cycles to the beginning of the list when it reaches the end.

The strategy is to use this feature to produce the desired non-periodic oscillations with a programmatically generated list of random frequencies. The elements of the duration list are set to be the multiplicative inverses of the elements of frequency list so that each frequency step lasts for one period. This oscillation is not only non-periodic but also random in that each cycle of the oscillation has a different period that is randomly chosen and therefore unrelated to other cycles before or after.

6.4.3 The Distribution of the Random Variable that Generates the Frequency List

For simplicity, the program written to generate the input list file uses a uniformly distributed pseudorandom variable to generate frequencies. Two parameters must be set. One is Δf , the width of the range of frequencies from which each random frequency is selected. The other is f_0 , the center of the range. The center was chosen to be $f_0=10$ kHz to be similar to the previous cathode experiment and actual Hall thruster oscillations. Reasonable values for Δf were

more challenging to set, but the values of $\Delta f = 3$ kHz and $\Delta f = 5$ kHz were ultimately selected for the experiment after performing simple simulations of a triggered average. The simulations of a triggered average that were used to make this selection are the subject of the remainder of this subsection.

Simulations of a triggered average of many instances of the type of random waveform discussed in Subsection 6.4.2 were done for several values of Δf . $Q = 10000$ instances of the random waveform were generated and averaged according to the process described in Subsection 4.3.4 to generate the simulated triggered average waveform. The combination of large Q and absence of noise ensures very little variation between different runs of the simulation, hence the triggered average waveform for any run of the simulation is essentially the expected value of the average waveform.

The range Δf was selected so that the decay of the simulated triggered average was qualitatively similar to the decay of the actual triggered average of TRLIF signal measured during a preliminary test of the TFLIF system with the X2 Hall thruster. Figure 6.8 shows the simulated triggered averages plotted together with the actual triggered average of TRLIF signal from a test with the X2 in August 2011.

To characterize the decay and frequency of the simulated triggered averages, a least squares fit was found to a decaying sinusoidal function of the form:

$$F(t) = (F_0 - F_u)e^{-kt} \cos(2\pi ft + \varphi) + F_u, \quad (6.1)$$

where the fit parameters are F_0 (initial value at $t = 0$), F_u (limiting value as $t \rightarrow \infty$), k (exponential decay rate), f (frequency), φ (phase offset). Figure 6.8 demonstrates that this functional form provides a good fit to the data, especially for the simulations. The mean absolute relative error of the least squares fit is 1.0%, 1.3%, and 2.6% for the $\Delta f = 3$ kHz simulation, $\Delta f = 5$ kHz simulation, and the X2 test, respectively. The fit is not as good for the X2 test, especially at later times, possibly due to the random noise inherently included in the LIF signal or possibly because the thruster's oscillations are not sinusoidal with uniformly distributed periods. The decay rates and frequencies of the fits are $k = 1.558$ kHz and $f = 9.933$ kHz for $\Delta f = 3$ kHz, $k = 4.602$ Hz and $f = 9.772$ kHz for $\Delta f = 5$ kHz, and $k = 6.477$ kHz and $f = 8.232$ kHz for the X2 measurement. Thus the simulated triggered averages with the selected values of Δf do indeed provide qualitatively similar behavior.

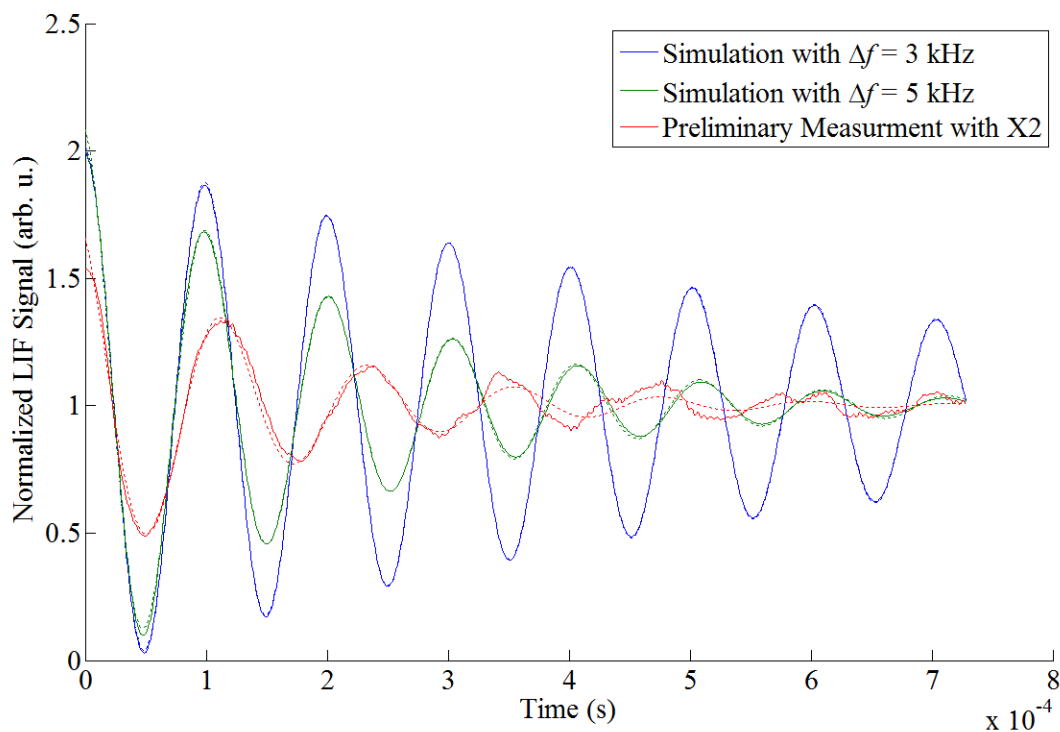


Figure 6.8. Plot of simulated triggered averages for $\Delta f = 3$ kHz (blue) and $\Delta f = 5$ kHz (green) together with the measured triggered average of TRLIF signal from a preliminary test with the X2. Least squares fits to the functional form in Equation (6.1) are also plotted in dotted lines.

6.4.4 Consistency Checks on the Properties of the Simulated Triggered Average

Analytical calculation of the expected triggered average waveform is difficult due to the iterated dependence on many random variables and is beyond the scope of this dissertation. Nonetheless, the strong least squares fits give reasonable confidence that the intuitive guess in Equation (6.1) is either close to or exactly the correct functional form. We can also make some simple calculations to check the consistency of the simulations with expected results.

It is interesting to note that the period of the average waveforms in Figure 6.8 is slightly longer than $1/f_0$, a result that is barely visible in Figure 6.8 but is evidenced by the frequencies slightly shorter than f_0 found by the least squares fits in Subsection 6.4.3. The expected period of the distribution can be calculated as:

$$\begin{aligned}
E[T] &= \int_{-\infty}^{\infty} T(f)P(f)df \\
&= \int_{f_0-\Delta f/2}^{f_0+\Delta f/2} \frac{1}{f} \frac{1}{\Delta f} df \\
&= \frac{1}{\Delta f} \left[\ln\left(f_0 + \frac{\Delta f}{2}\right) - \ln\left(f_0 - \frac{\Delta f}{2}\right) \right] \\
&= \frac{1}{\Delta f} \ln\left(\frac{f_0 + \frac{\Delta f}{2}}{f_0 - \frac{\Delta f}{2}}\right),
\end{aligned} \tag{6.2}$$

where $T(f)$ is the period, $P(f)$ is the probability density function of the uniform random variable, which has a value of $P(f) = 1/\Delta f$ for $f_0 - \Delta f/2 < f < f_0 + \Delta f/2$ and zero otherwise. This result gives an expected period of 1.0076×10^{-4} s for $\Delta f = 3$ kHz and 1.0217×10^{-4} s for $\Delta f = 5$ kHz, which compare well with the periods from the least squares fits of 1.0067×10^{-4} s for $\Delta f = 3$ kHz and 1.0233×10^{-4} s for $\Delta f = 5$ kHz.

A plot of the phase difference of a set of simulated random waveforms relative to a sinusoid at the central frequency $f_0 = 10$ kHz is shown in Figure 6.9. For a given frequency f , phase relative to f_0 drifts linearly in time at a rate of $2\pi(f - f_0)$. A new random frequency is selected at the end of a period, and therefore the drift rate changes to some other constant for the duration of the next period, leading to the random piecewise linear functions found in the plot. There is a clear trend toward negative phase drift that might be naively interpreted to indicate a bias in the code to select frequencies lower than f_0 or perhaps some other issue.

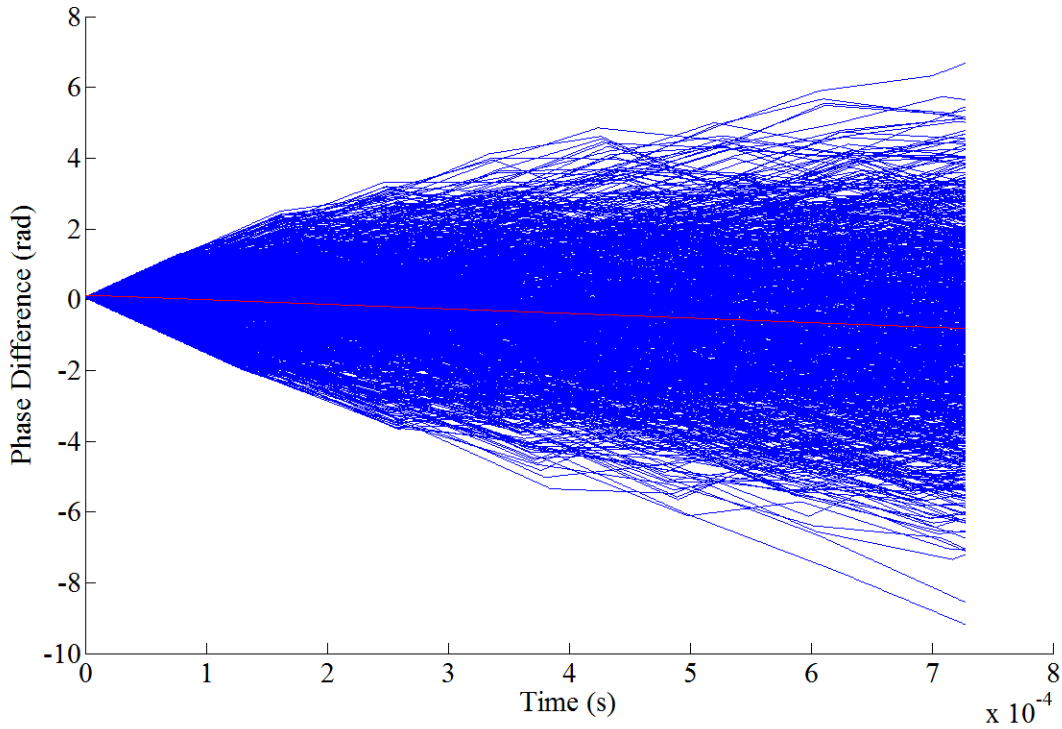


Figure 6.9. Plot of the phase difference of 1000 instances of the random waveform relative to the phase of a sinusoid at a frequency f_0 (blue lines) for the $\Delta f = 5$ kHz case. The linear fit representing the average phase drift of the random waveforms from the sinusoid at f_0 is also plotted in red.

On the contrary, negative average phase drift is expected even for random frequencies selected from a uniformly distributed range centered on f_0 . Intuitively, the reason for this is that slow frequencies ($f < f_0$) will have longer periods and therefore spend more time drifting negatively than fast periods ($f > f_0$) will spend drifting positively during a single period. The phase drift per period as a function of random frequency is simply the drift rate multiplied by the period $T = 1/f$:

$$\Delta\varphi = 2\pi(f - f_0)T = 2\pi(f - f_0)/f. \quad (6.3)$$

Then the expected value of $\Delta\varphi$ can be calculated similarly to the expected period since it is also a function of the random variable f :

$$\begin{aligned}
E[\Delta\varphi] &= \int_{-\infty}^{\infty} \Delta\varphi(f)P(f)df \\
&= \int_{f_0-\Delta f/2}^{f_0+\Delta f/2} 2\pi \frac{f-f_0}{f} \frac{1}{\Delta f} df \\
&= 2\pi \left(1 + \frac{f_0}{\Delta f} \left[\ln\left(f_0 - \frac{\Delta f}{2}\right) - \ln\left(f_0 + \frac{\Delta f}{2}\right) \right] \right) \\
&= 2\pi \left(1 + \frac{f_0}{\Delta f} \ln\left(\frac{\frac{f_0}{\Delta f} - \frac{1}{2}}{\frac{f_0}{\Delta f} + \frac{1}{2}}\right) \right).
\end{aligned} \tag{6.4}$$

The result is always negative for values of the ratio in the range $1/2 \leq f_0/\Delta f \leq \infty$ (all of the physically possible values) and it monotonically increases in this range approaching zero as $f_0/\Delta f$ approaches infinity.

This expression gives an expected phase drift per period of $E[\Delta\varphi] = -0.048$ rad for $\Delta f = 3$ kHz and $E[\Delta\varphi] = -0.136$ rad for $\Delta f = 5$ kHz. Now, the slope of the linear fit of the simulated phase drift is -454 rad/s for $\Delta f = 3$ kHz and -1293 rad/s for $\Delta f = 5$ kHz, leading to a drift per expected period of 0.046 rad for $\Delta f = 3$ kHz and 0.130 for $\Delta f = 5$ kHz, in reasonable agreement with the analytical result.

6.4.5 Advantages and Limitations of List Mode

Many function generators have an arbitrary waveform feature that is typically meant to specify an arbitrary cycle of a periodic waveform. It could be used to generate many random cycles with arbitrary waveforms, but would be limited by the memory available since all points in the waveform are explicitly stored in memory. For example, the Agilent 33220A supports arbitrary waveforms up to a maximum of 65536 samples. With a bare minimum of 10 samples per oscillation period, it could support at most 6553 independent cycles, corresponding to about 0.65 s at an approximate cycle period of 10^{-4} s. Then this waveform would have to be repeated about 100 times during an acquisition of 60 s, and the output would not be smooth given so few samples per period.

On the other hand, List Mode's advantage is that many more random cycles can be smoothly generated since only the frequency and duration lists are stored in memory, not all points. The maximum list length that would be accepted in List Mode was found to be 99,999. Since the average duration is about 10^{-4} s, this leads to a random waveform for a total time of

about 10 s before the list recycles. The need for averaging requires the experiment to take 60 s of data per wavelength, hence the random waveform is actually repeated about 6 times during each 60-s data set. This repetition somewhat reduces the randomness of the waveform, but it repeats over a time scale much longer than the chunk length used in averaging. The fact remains that any particular cycle is unrelated to tens of thousands of cycles before and after.

One limitation is that the function generator accepts values for frequency and duration only within a certain discrete set. The random values of frequency and duration in the list are not necessarily members of the discrete set of possible values for the function generator, in which case the function generator automatically rounds the frequency and duration values to the closest values possible for the function generator. When this happens, the duration no longer corresponds to exactly one period of the frequency.

The results is that each frequency step lasts for approximately but not exactly one period, and therefore the phase at which the frequency changes tends to drift over time. This effect is not a problem for the experiment, and in fact it tends to further randomize the signal despite the repeating sequence of frequencies. It is of note because the statement above that the oscillation is sinusoidal with period changing at each cycle is approximately but not exactly true. In a case where that exact waveform is necessary, it may be possible to circumvent the issue by randomly selecting only pairs of frequency and duration that are both allowed by the function generator. That complication in the code was not necessary for our purposes.

6.4.6 Results

A comparison between the TFLIF and TALIF signals for $\Delta f = 5$ kHz is shown in Figure 6.10. Average signal normalization is used. 18 wavelengths were probed. The input discharge current trace was chosen to begin at a trigger so that the results would be in phase at least initially.

As predicted for this random oscillation, there is a distinct nonphysical decay in the signal for the triggered average but not for the transfer function average. The transfer function average follows the frequency of the discharge current and the amplitude tends to be lower for higher frequency periods, which is also expected since the frequency range used is close to the amplifier's 20 kHz bandwidth.

These expected features lend credence to the result of transfer function average and the comparison clearly shows that the transfer function average does not fail as the triggered average

does. Unfortunately, this comparison alone cannot confirm the accuracy of the transfer function result as it did for the periodic oscillation since the triggered average fails in this case. Hence there is a need for a benchmark that can validate the transfer function average with nonperiodic oscillations.

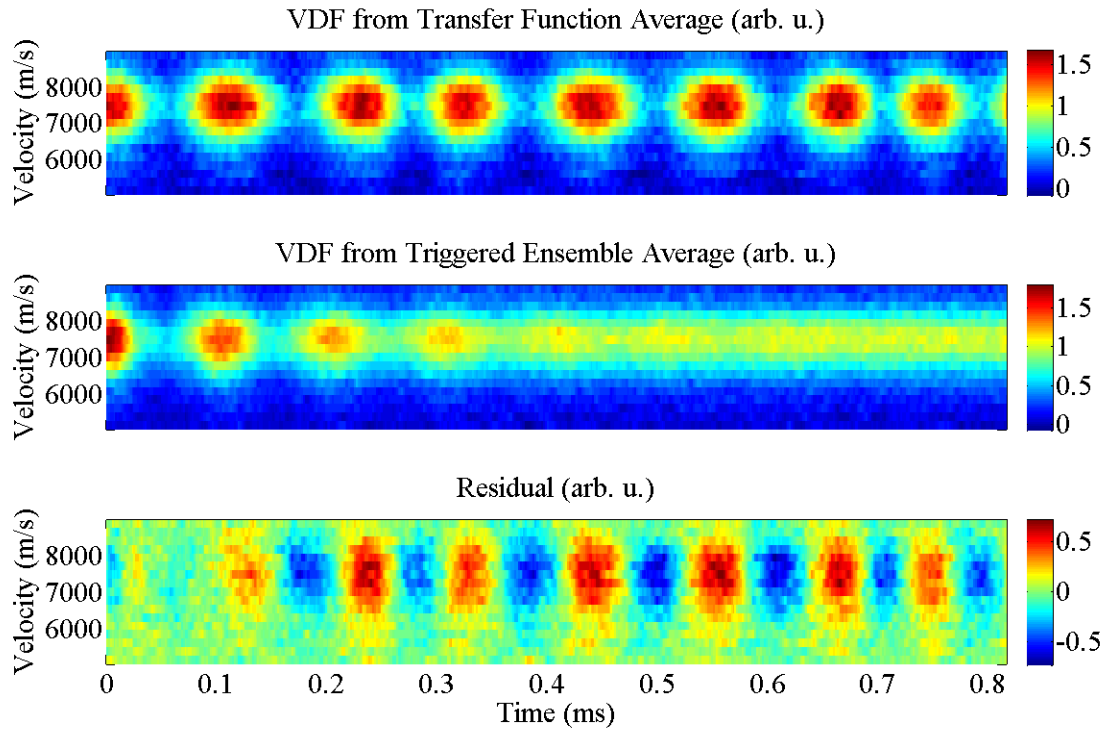


Figure 6.10. A Comparison between TRLIF signal from the transfer function average (top) and triggered average (middle) with the residual on the bottom. The input to the transfer function is an actual segment of discharge current beginning at a trigger so that they are initially in phase.

We use the benchmark proposed in Section 4.3.5. In short, Section 4.3.5 derived, assuming a linear relationship between discharge current and LIF signal, that the transfer function reproduces the triggered average LIF signal if the triggered average discharge current is used as input to the transfer function. Therefore, we validate the transfer function by comparing the actual triggered average LIF signal to the reproduction of that signal calculated via the transfer function.

Figure 6.11 is an example of this comparison. The same dataset is used as in Figure 6.10. The transfer function itself is also the same; the only difference is that the triggered average of discharge current traces is used as the input to the transfer function. The RMS residual is 0.0788, about 8% of the average signal level at the peak. The residual is largely random noise and there

is good agreement between the two signals, which is evidence that the transfer function analysis works and the assumptions made are reasonable for this physical system.

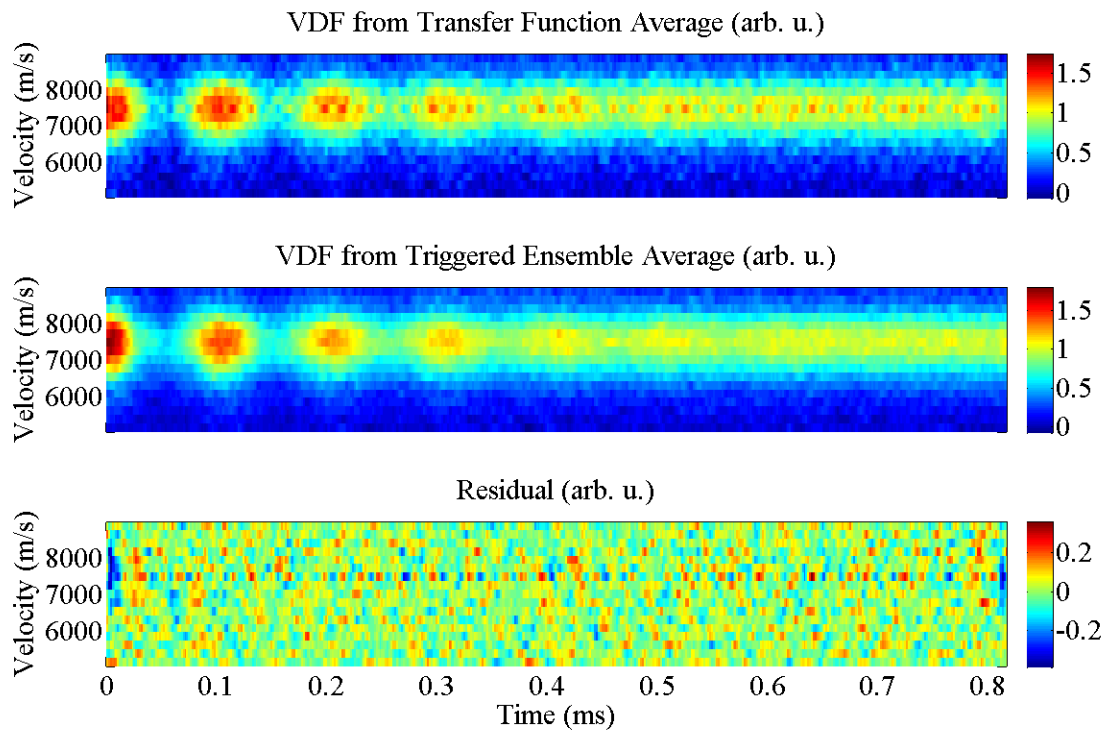


Figure 6.11. A Comparison between LIF signal from the transfer function average (top) and triggered average (middle) with the residual on the bottom. The input to the transfer function is the triggered average of discharge current traces. In this case, the TFLIF signal does reproduce the TALIF signal as expected.

Unfortunately, there are two types of minor systematic error to note: (1) incorrect amplitude at the center of the distribution near the edges of the window, and (2) high frequency noise at a single laser wavelength corresponding to a velocity group near the center of the VDF. Fortunately, Appendix A shows that both of these errors are due to the choice of window type and chunk size used in the analysis.

A rectangular window was used for this comparison since it allows the whole chunk to be viewed unobscured by the window function. A Tukey window is used for most other analyses in this dissertation, which reduces both random and systematic error, but for this comparison it is important to plot the whole time available in the chunk since we want to capture the decay of the triggered average and demonstrate that the transfer function reproduces it. The same comparison with the Tukey window gives very good agreement, but the plot is less intuitively convincing

because much of the decay has already occurred before the central region of the chunk that the Tukey window allows to be plotted (see Appendix A).

6.5 Summary of Arguments Validating TFLIF

In summary, the results of the campaign suggest that the TFLIF system accurately recovers the sinusoidal and even the randomly varying oscillation. There is no decay or other obvious artifact in the TFLIF signal for the random oscillation synthesized with the actual discharge current input (top frame of Figure 6.10). The TFLIF signal shows two properties that are expected under the circumstances of the hollow cathode test. Firstly, it follows the frequency of the discharge current oscillation. Secondly, the amplitude varies as a function of frequency due to operating near the -3 dB frequency of the power amplifiers driving the discharge current.

In addition to the evidence above, several quantitative benchmarks were verified. These are specific cases where the TFLIF signal theoretically converges to another signal. For the first benchmark, Figure 6.4 verifies that the TFLIF, TALIF, and lock-in amplifier measurements of the time-averaged LIF profile all agree. The test with sinusoidal oscillations shows that the TFLIF signal (synthesized with an actual discharge current trace) converges to the same result as the TALIF signal (Figure 6.5). This property is expected when the assumptions of both averaging techniques are satisfied since in that case they both converge to the actual TRLIF signal (see Subsection 4.3.4). Recall that for TFLIF the assumption is a linear system relating input to output, while for TALIF the assumption is that oscillations are periodically repeated after each trigger.

For the randomly varying oscillation of Section 6.4, Figure 6.10 shows that the TFLIF and TALIF signals no longer converge together because the triggered average fails due to the nonperiodic oscillations. In this case, the general-purpose benchmark of Subsection 4.3.5 can be used. Figure 6.11 verifies that the average transfer function reproduces the TALIF signal as theoretically expected when the characteristic output is calculated for the input signal of the triggered average discharge current (see Subsection 4.3.5). N.B. the signal of Figure 6.11 is nonphysical since the triggered average fails. Nonetheless, it shows that the average transfer function can reproduce the signal as expected, and therefore the average transfer function has converged close enough to the actual transfer function to give an accurate characteristic output signal. Though the comparison of Figure 6.11 is not physically meaningful, the verification of the expected property gives some evidence that the same transfer function will also give accurate

results when calculating the characteristic output signal for a physically measured discharge current trace, whose output will be physically meaningful.

These results are ultimately a validation only for the hollow cathode plasma source, though the demonstration that the TFLIF technique works in one case suggests it may work with other plasma sources. Results for other sources and conditions not explicitly tested here (such as the more complicated plasma dynamics of the Hall thruster) will need to be validated separately since the assumption of a linear system relating input and output signals may not always hold. A comprehensive validation will include three components. First, a similar series of quantitative benchmarks as above. These benchmarks provide limited evidence. They do not provide a deductive logical argument, but merely check the consistency of the signal with expected properties. It may be possible for a signal to have significant error while still satisfying the benchmarks. A second part of validation should be a careful consideration of the implications of the results and whether they are physically reasonable. Thirdly, validation should also include a consideration of how the TFLIF measurements compare with other high-speed measurements (e.g. high-speed Langmuir probe, high-speed camera, and other TRLIF techniques) and with theory and simulation.

Chapter 7

Initial Hall Thruster Campaign

If there is no other Use discover'd of Electricity, this, however, is something considerable, that it may help to make a vain Man humble.

– Benjamin Franklin (Letter to Peter Collinson, 14 Aug 1747)

Having covered the basic theory and implementation of the new technique as well as preliminary validation experiments with a hollow cathode, the capstone of this dissertation on the development of the TFLIF technique is an initial campaign with a Hall thruster. This chapter has two goals. The first is to show that the technique is capable of accurately recovering TRLIF signal from a Hall thruster. In contrast to the previous chapter, where experiments were designed solely for validation, the second goal of this chapter is to gather new Hall thruster measurements that are important in their own right. The first practical application of TFLIF on a Hall thruster confirms major changes in ion dynamics depending on the magnetic field setting of the thruster. This chapter demonstrates signal recovery in both approximately periodic and nonperiodic thruster operating modes; a first for time-resolved LIF measurements in Hall thrusters.

The key assumption of the TFLIF technique is that the thruster acts as a time invariant linear system with some transfer function mapping the input signal (such as discharge current or floating probe voltage) to the LIF signal as the system output. Due to the time invariance, the transfer function itself is assumed constant while the thruster operates at single operating condition in equilibrium, but the individual oscillations that go into calculating the ensemble of empirical transfer function estimators need not be periodic or repeatable. Not only may amplitude and period vary randomly, as shown in the cathode validation results of the previous chapter, but a slow drift in the Hall thruster oscillating frequency will not affect the measurement as long as the transfer function itself is constant.

7.1 Experimental Configuration

7.1.1 Experimental Setup in the Vacuum Chamber

A diagram of the experimental setup inside the Large Vacuum Test Facility (LVTF) is shown in Figure 7.1. The laser, propagating along the thruster axis to measure axial velocity, is focused to a 1 mm spot near the thruster exit plane. A 75-mm-diameter lens with 85-mm focal length images light collected from the interrogation volume onto a 1-mm optical fiber with unity magnification. The collection optics are contained in a lens tube (visible in Figure 7.2 to the right of the thruster) with a sacrificial glass window with an antireflective coating. Light was collected 30 degrees from the exit plane, defining a small interrogation volume about 1 mm in all dimensions. All points interrogated are along the discharge channel centerline at the thruster's 3 O'clock position. The laser is radially polarized to excite primarily π transitions in the thruster's radial magnetic field, reducing Zeeman splitting [100]. The thruster itself is mounted on x-y motion stages to move the stationary interrogation volume relative to the thruster so that the sensitive optics need not be moved. LIF is generally considered a non-intrusive diagnostic, or at least less intrusive than probes, and as such there are no pieces of the experimental setup capable of disturbing the discharge more than normal facility effects.

The H6 processes much more power than the cathode used in the previous campaign, causing laser alignment issues apparently from thermal drift as structures heated. Both optics and their structures are protected from ion bombardment behind graphite shields and glass windows. The shielding and window are supported by independent structures and entirely block the line-of-sight from the thruster and ion beam to the injection optics and support structure. In addition, the polar and azimuthal angles of the injection optics were controlled remotely by motors to correct drift on the fly and maximize LIF signal. Correction was routinely performed approximately hourly by maximizing the laser reflection off of a 1-mm diameter pin fixed a few centimeters aside the thruster (visible just to the right of the thruster in Figure 7.2). The drift was slow enough that alignment drift is an insignificant source of uncertainty for the LIF signal amplitude at different wavelengths in the same VDF. The measured velocity distributions at each spatial point or thruster operating condition are normalized independently and laser alignment may slightly change, therefore signal level is not comparable between VDFs.

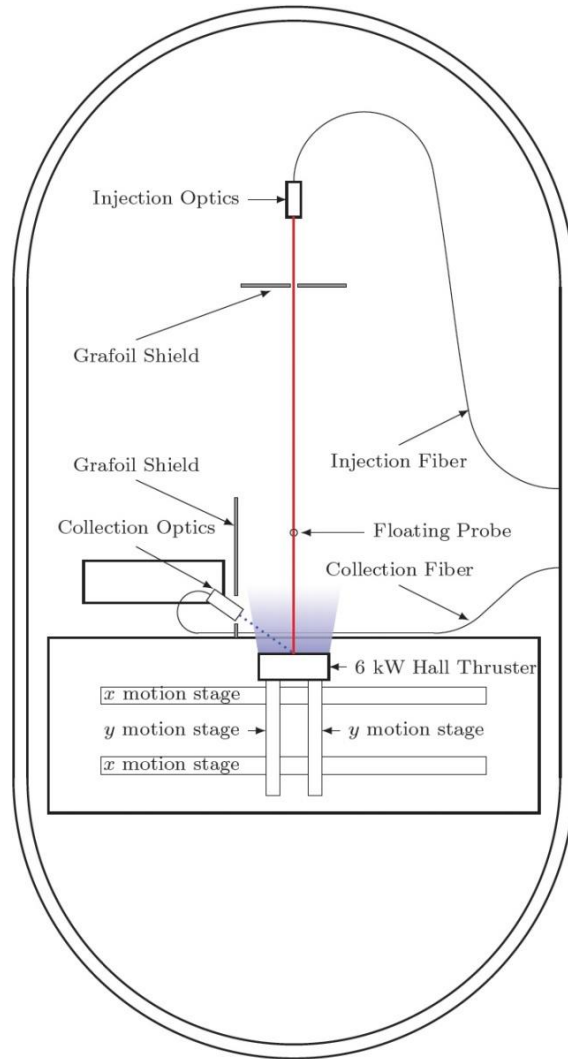


Figure 7.1. Diagram of the experimental setup inside the Large Vacuum Test Facility. Reproduced from Ref. [101].

The outer magnetic pole piece of the H6 was found to be visibly warped during the setup for this campaign (see Figure 7.3). The originally flat pole piece bows with a maximum deflection of about 1 mm near the 11 O'clock position. The warping may be due to an excessive heat load on the pole piece during a previous campaign, but the exact origin is unknown. The warping is small enough that the magnetic field and therefore operation of the thruster is probably not significantly altered.

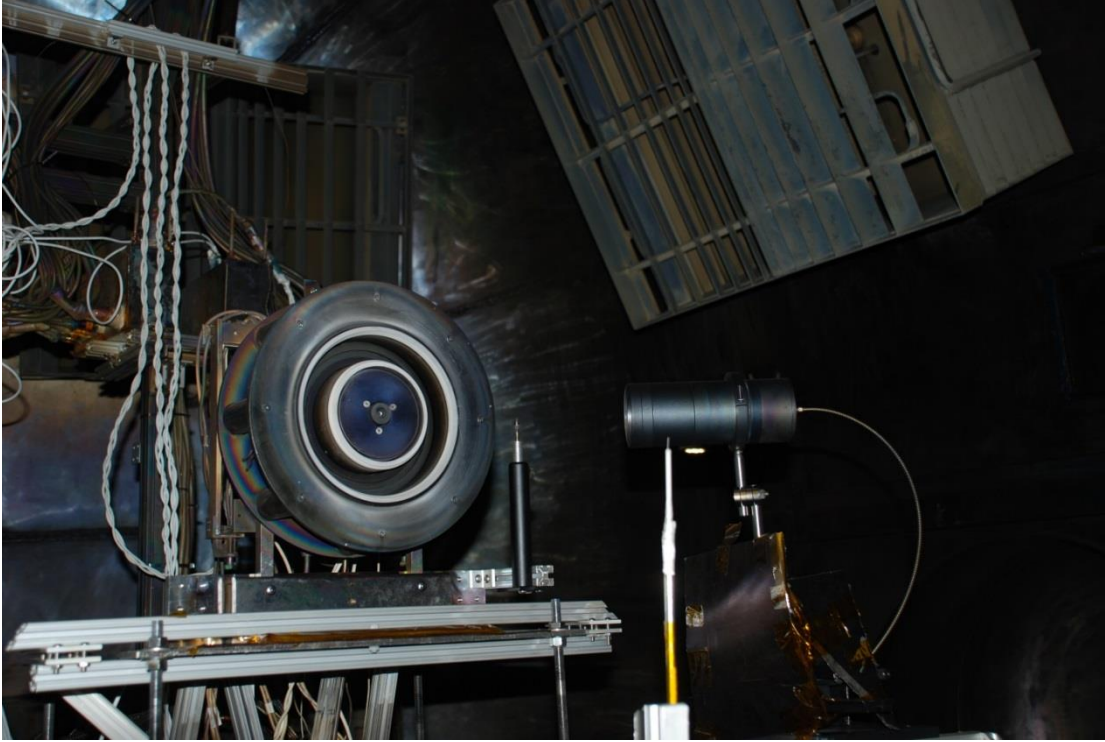


Figure 7.2. The H6, alignment pin, floating probe, and collection optics photographed in position for the campaign.

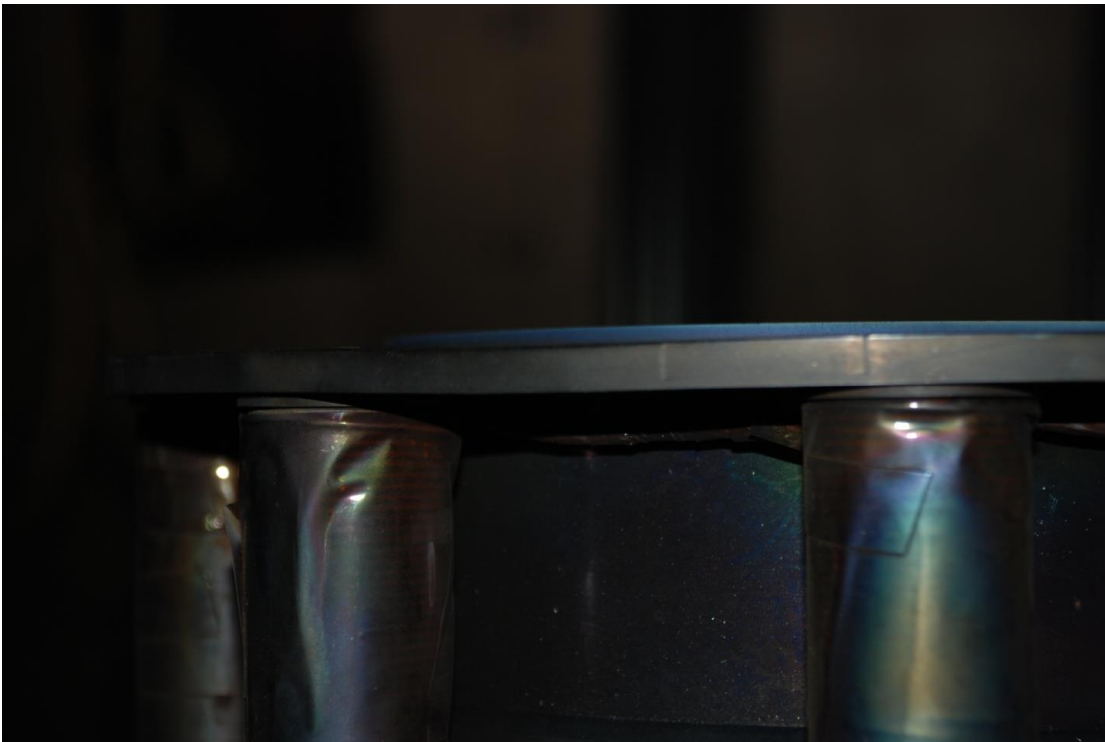


Figure 7.3. Photograph of the warped outer pole piece near the 11 O'clock position of the H6.

7.1.2 Auxiliary Measurements

A floating probe placed five thruster diameters downstream along the 6 O'clock channel centerline (supported by a white, vertical alumina tube visible in Figure 7.2) is an input to the transfer function that can be used alternatively to the discharge current. Two Alazartech ATS9462 cards digitize all three signals simultaneously to enable comparison between results with the two simultaneous input signals.

High-speed video of the thruster with the Photron FASTCAM SA5 nearly head-on was captured in this campaign to compare with TFLIF data. The high-speed image analysis technique was discussed in Subsubsection 2.5.2 and described in detail by McDonald [8] and Sekerak [10]. For this campaign, the camera captured a resolution of 256×256 at a frame rate of 87.5 kHz for a total of 2^{14} frames, or about 0.19 s. To observe the optical emission at the same time as the other signals, the camera was triggered simultaneously with the digitizers measuring PMT voltage, discharge current, and floating probe potential.

Alicat flow controllers with a maximum of 50 sccm (serial number 85697) and 400 sccm (serial number 57639) controlled cathode and anode flow, respectively. Both were calibrated using a Bios Definer 220 in July 2014, a few months before this campaign. The actual flow rate at five different settings were measured and a linear fit gives a function that maps the actual flow rate to the setting of the flow controller. Figure 7.4 gives the resulting calibration curves. R^2 values near unity show that they are both very linear over the whole range tested. Note that the gain changed by less than 1% for the anode and about 3% for the cathode from the previous calibration in February 2013, while the offset changed from the previous calibration values by about 25% (for anode) and 64% (for cathode). The large relative change of the offset is insignificant given that it is a fraction of 1 sccm.

The Definer was operated in continuous mode averaging over 10 samples per measurement. Measurements were discarded until the back pressure and flow measurements equilibrated so that each measurement varied only by a small fraction of 1 sccm. A slow drift was observed in the calibration measurement even after reaching equilibrium in backpressure and flow rate measurements. An example of the drift observed is given in Figure 7.5 for the 50 sccm setting of the anode controller. The drift made defining a value for a given flow rate difficult and time consuming. The drift curves at all settings are similar in that there is a relatively long and linear drift at first, followed by a slower drift. The flow rate value at each

flow rate setting was defined by excluding the faster part of the drift, which apparently occurs before the system is truly in equilibrium, and then averaging all measurements thereafter. In the example case of the 50 sccm setting of the anode controller, the measurement value was defined as the average of measurements 13 and after. The origin of the drift is unclear, but it does represent a small uncertainty in the measurement of less than 1 sccm.

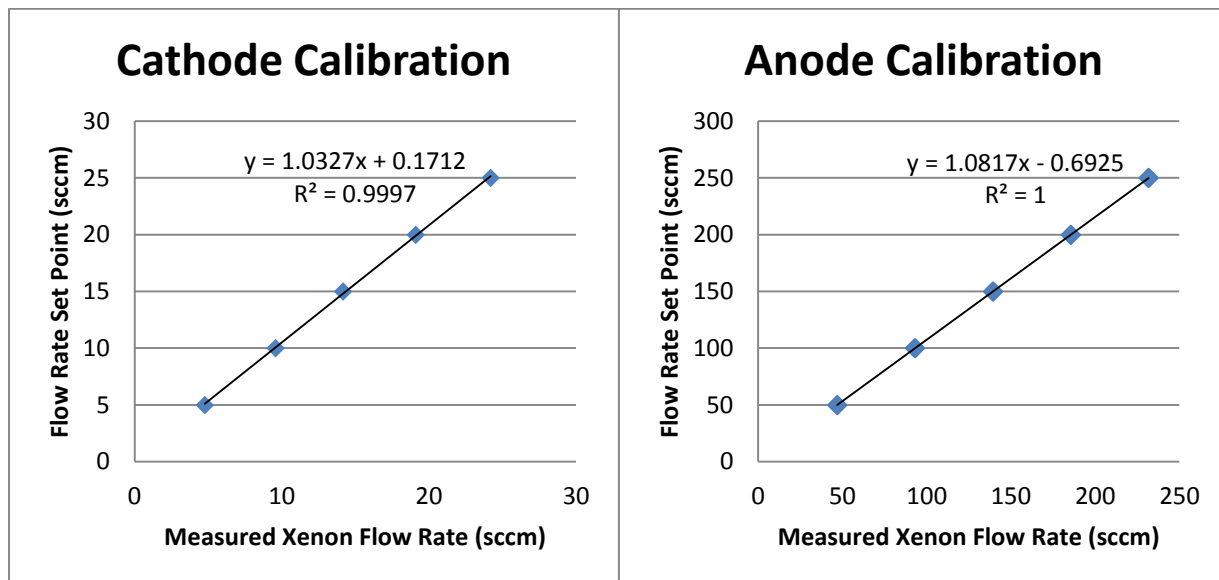


Figure 7.4. Calibration curves for the cathode (left) and anode (right) flow controllers.

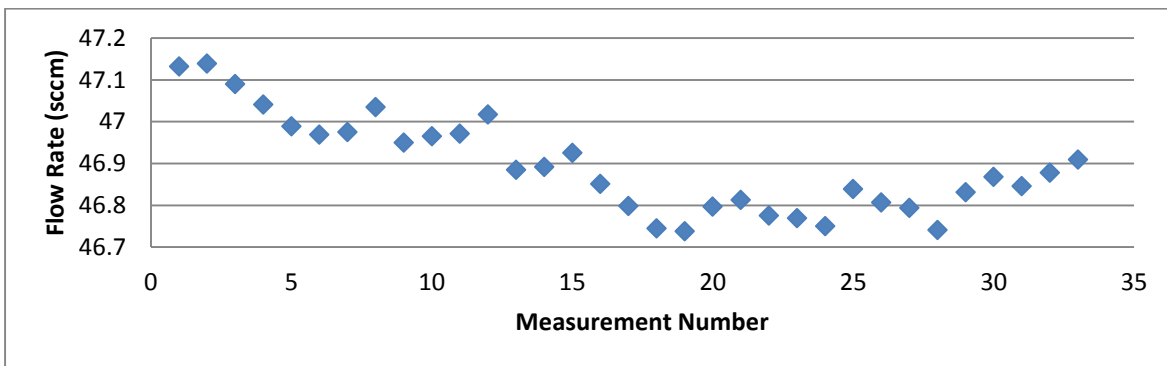


Figure 7.5. Example of flow rate measurement drift observed after equilibrium in back pressure at a constant flow rate setting of 50 sccm on the anode controller.

7.1.3 Test Matrix

A discharge voltage of 150 V and an anode flow rate of 10 mg/s were used to maximize the LIF signal-to-noise ratio (SNR). The SNR was much lower at discharge voltages of 150 V and 300 V than previously reported with the H6 at this facility [45]. The cause is currently

unknown but will be investigated to allow measurement at the nominal 300 V condition. The cathode flow fraction was 7%.

The loss of laser alignment due to the unstable table supporting the H6, motion stages, and collection optics consumed much of the time allotted for this campaign. The problem was temporarily solved by locking the table's 11 independent sections together with 80/20, tying each section to various mounting points in the chamber with wire rope, and moving the collection optics to an 80/20 structure off of the table. The table was removed from service after this campaign and a new support structure was designed and built for use with the motion stages.

Three axial locations along the channel centerline were interrogated to investigate how the ion VDF evolves as the ions are accelerated. The VDF is also measured at the exit plane under three magnetic field settings from near the lowest field that gives stable operation up to nearly saturating the magnetic circuit. Table 7.1 summarizes the different conditions interrogated. Many spatial locations and a more complete magnetic field sweep were originally planned. The campaign was restricted to one month of chamber time, and the issues with the table took enough time to necessitate paring down the plan for this campaign to the smallest set of points that could still yield interesting results. The data presented here were all taken in two days near the end of March 2015.

Table 7.1. Test matrix of conditions where TFLIF signal was captured. The campaign was entirely at the H6 150-V, 10-mg/s operating condition with all LIF data taken along the channel centerline at 3 O'clock. Green cells denote data was taken and presented. Analysis failed to yield reasonable results for data taken at the exit plane with nominal field (red cell) for unknown reasons.

		Z (mm)		
		- 4	0	15
B_r/B_r^*	0.5238			
	1			
	1.48			

7.2 Results and Discussion

7.2.1 Evolution of the Velocity Distribution Between Two Axial Positions

Time-resolved ion VDFs were measured at two locations at nominal magnetic field (inner magnet current 3.5 A). These points demonstrate data collection in both the relatively hot and dense plasma in the discharge channel and the cooler and rarer plasma in the plume. The point in the discharge is 4 mm upstream of the exit plane ($z = -4$ mm), and the point in the plume is 15 mm downstream ($z = 15$ mm). The points were chosen based on electron density and temperature maps at 300 V from Reid [69]. Density and temperature maps were not reported at 150 V, but maps of electric field at 150 V in Appendix C of the reference show that the plasma properties likely have a similar spatial trend.

The measured ion velocity distribution at $z = -4$ mm is shown in Figure 7.6. There is an approximately periodic oscillation with a high density population of ions forming at a minimum in mean velocity near 3.5 km/s. The mean velocity then increases to a maximum of about 5 km/s. After reaching the maximum velocity, the population density and distribution spread suddenly fall and the mean velocity declines back to the minimum at approximately the same rate as the rise in velocity. The oscillation then begins again, leading to a shape similar to a triangle wave. There is a long low energy tail extending almost to 0 m/s that only appears when the bulk velocity is near its minimum.

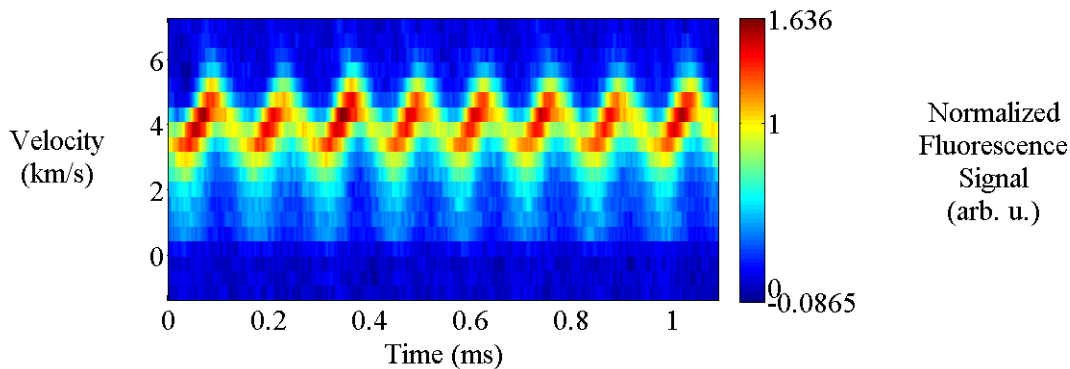


Figure 7.6. The Ion velocity distribution at $z = -4$ mm for the 150-V, 10-mg/s operating condition of the H6 with nominal magnetic field shows an oscillation in central velocity, distribution spread, and relative metastable density.

The measured ion velocity distribution at $z = 15$ mm is shown in Figure 7.7. The LIF signal was much weaker at this point, which is apparent in the lower signal-to-noise ratio, though

the maximum in the time-averaged signal is normalized to 1. The basic features are similar to the $z = -4$ mm point except the distribution is accelerated, as expected. A similar oscillation in mean velocity occurs between about 10 and 12 km/s at $z = 15$ mm. Density similarly falls as the mean velocity begins to decline, but the rate of decline increases slightly toward the minimum in mean velocity, leaving a small kink in the downward part of the plot. The maximum in bulk distribution spread occurs at the minimum of mean velocity. The low velocity tail is less pronounced at this location and is largely obscured by the noise.

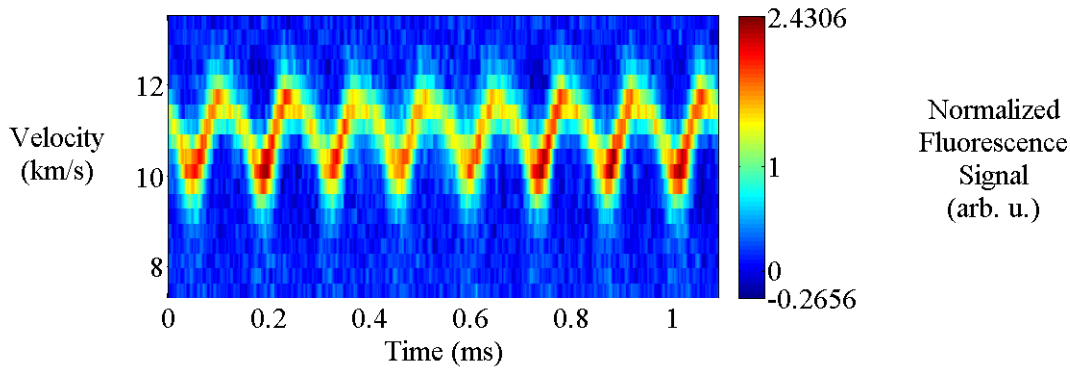


Figure 7.7. Ion velocity distribution as a function of time at $z = 15$ mm for the 150-V, 10-mg/s operating condition of the H6.

The distribution is significantly narrower downstream. The change in the distribution spread can be partially explained by kinematic compression or velocity bunching, an effect whereby a velocity spread narrows due to an acceleration [102],[103]. Intuitively, it occurs because fast ions spend less time in the accelerating potential than slow ions, and therefore receive a smaller increment in velocity, hence the velocity spread between them is reduced.

Bunching can make a prediction of the distribution FWHM at $z = 15$ mm based on the FWHM at $z = -4$ mm and the potential drop observed between $z = -4$ mm and $z = 15$ mm. For simplicity, we assume that the distribution at $z = 15$ mm will only depend on the accelerated ions from the $z = -4$ mm distribution at the same time. In reality, the ions at $z = 15$ mm at a certain time will have come from $z = -4$ mm at slightly different times (and some even may have been born downstream of $z = -4$ mm), but the simplifying assumption is reasonable since even slow ions with an average speed of 3 km/s will travel the 19 mm distance in about $6 \mu\text{s}$, a time in which the VDF changes little and that is near the time resolution of this dataset.

The ratio of the measured width ($FWHM_{meas}$) to the predicted width ($FWHM_{theory}$), plotted in Figure 7.8, oscillates between about 1 (about the same as predicted) and 3 (3 times as wide as

predicted). The plot is very noisy due to the uncertainty in measuring FWHM in both distributions and the potential drop between the two points. The oscillation in the ratio of $FWHM_{meas}$ to $FWHM_{theory}$ is strongly correlated with but slightly out of phase with the discharge current oscillation. This behavior may be evidence of an ionization zone oscillating in axial position. When the ratio is greater than 1, there may be significant ionization downstream of $z = -4$ mm, leading to a wider distribution than predicted by bunching alone calculated from the distribution at $z = -4$ mm. When the ratio is near 1, there may be little ionization downstream of $z = -4$ mm, and therefore distribution spread at $z = 15$ mm downstream is well explained by bunching. The possibility of an oscillation in the position of the ionization zone is a hypothesis for which evidence has been observed in Hall thrusters (e.g. Mazouffre[79],[81]), but a more complete data set is needed to confirm it in the H6 at this operating condition.

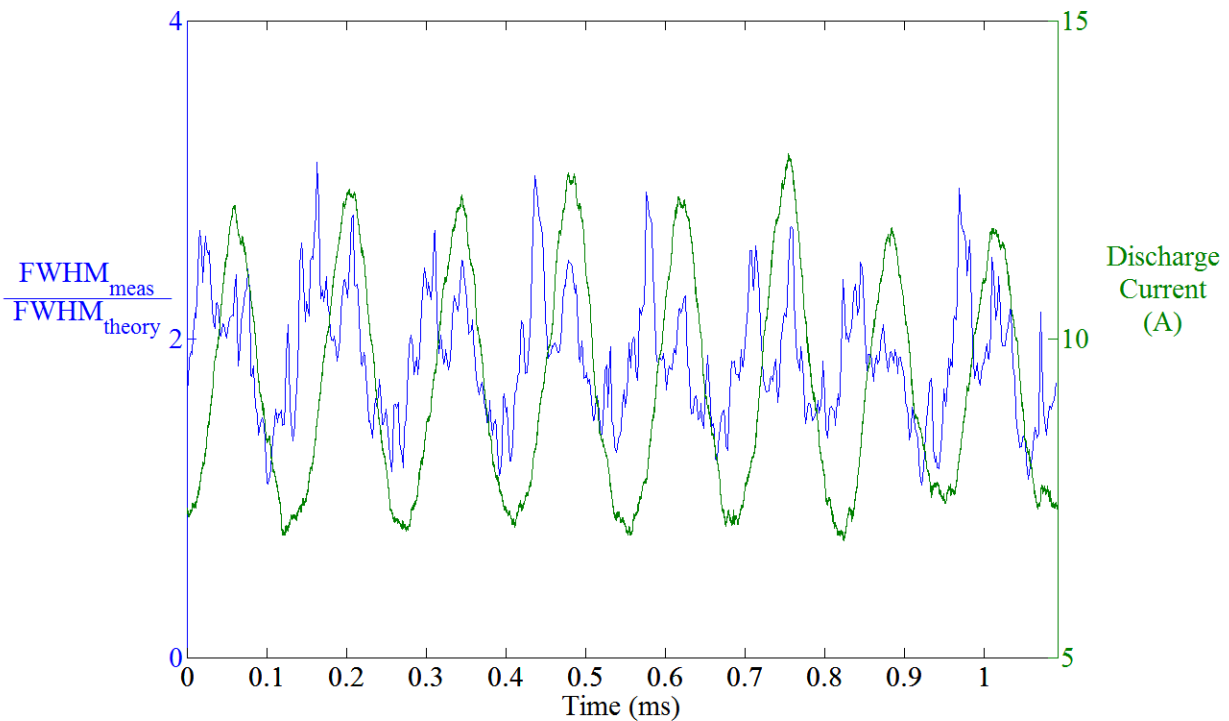


Figure 7.8. Ratio of measured ion velocity distribution FWHM at $z = 15$ mm to that predicted by bunching given the ion velocity distribution at $z = -4$ mm. It oscillates between approximately 1 and 3, highly correlated with discharge current oscillations, and possibly indicates axial motion of the ionization zone.

7.2.2 The Effect of Varying Magnetic Field

The thruster's magnetic field was varied to observe changes such as the thruster operating mode transitions observed with the H6 by Sekerak [10]. This reference also briefly presents a least squares fit to predict magnetic field magnitude from Infolytica MagNet simulations. The magnetic field magnitude, normalized by the nominal magnetic field magnitude B_r^* , is given as a function of inner magnet current in amperes I_{IM} by:

$$B_r / B_r^* = -0.0105I_{IM}^2 + 0.3343I_{IM} - 0.0444 . \quad (7.1)$$

This formula assumes a constant ratio of inner magnet current to outer magnet current of 1.12. This ratio was maintained throughout this campaign to investigate the effects of changing the magnetic field magnitude while keeping a constant field topology. A sweep of magnetic field to map out thruster operating modes at many field magnitudes was planned, but time allowed for only three magnetic field points to be measured. Though the three magnetic field settings are widely spread, care was taken to ensure the ratio of inner magnet current to outer magnet current was approximately kept constant while changing settings to avoid changing the magnetic field shape and operation of the thruster unnecessarily.

7.2.2.1 Auxiliary Measurements on the Thruster Operating Mode

Measurements of discharge current captured by a high-speed current probe (the same probe used for the transfer function input) and light emission captured by the Photron FASTCAM SA5 high-speed camera help assess the thruster operating mode and mode transitions. Global oscillations are contained in the discharge current and the average over all pixels in each frame of the FASTCAM video, while the azimuthally local oscillations are captured by the power spectral density analysis of the FASTCAM video (see Subsections 2.6.1 and 2.5). The mode transition criteria determined for the H6 for operating conditions near a discharge voltage of 300 V and a flow rate of 20 mg/s are discussed in Subsubsection 2.6.1.2.

Figure 7.9 presents the power spectral density of the FASTCAM intensity at the three magnetic field settings tested. The low-field condition has by far the strongest global mode oscillation with sharp peaks and four harmonics visible up to the Nyquist frequency. The spoke modes have peaks at the same frequencies as the global mode, an effect also observed at low-field magnitude with a discharge voltage of 300 V and anode flow rate of 19.5 mg/s [10]. For the measurement at the low-field setting with an anode potential of 150 V and an anode flow rate of

10 mg/s, the global mode dominated and there were no visible azimuthal perturbations in the high-speed video. Thus, the relatively strong $m > 1$ peaks are likely smearing from the $m = 0$ peaks, as described by McDonald [8]. Turbulent azimuthal perturbations were visually present at the nominal field but the (weaker) global oscillation continued to dominate. Oscillations were weaker and less coherent since the maxima at all peaks dropped, peaks were broader, and many of the higher harmonics were no longer resolved above the noise floor. The high-field condition had the weakest and broadest global mode peak with no harmonics, indicating the least coherence of the three settings. The spoke mode peaks had mostly vanished, but the power spectral density of the noise floor is higher than at the other field magnitudes. The strongest azimuthal perturbations were observed at this condition, but they were highly random and not coherent spokes due to the spectra similar to white noise.

In general, the peaks observed here tend to decrease in frequency with increasing magnetic field, in agreement with Sekerek's results. The transition from relatively coherent global oscillations to relatively incoherent azimuthal oscillations with increasing field is reminiscent of Sekerek's magnetic field sweeps (cf. Figures 17 and 18 in reference [10]). Those figures, however, show that spoke mode peaks intensified with increasing field, not the effect observed here.

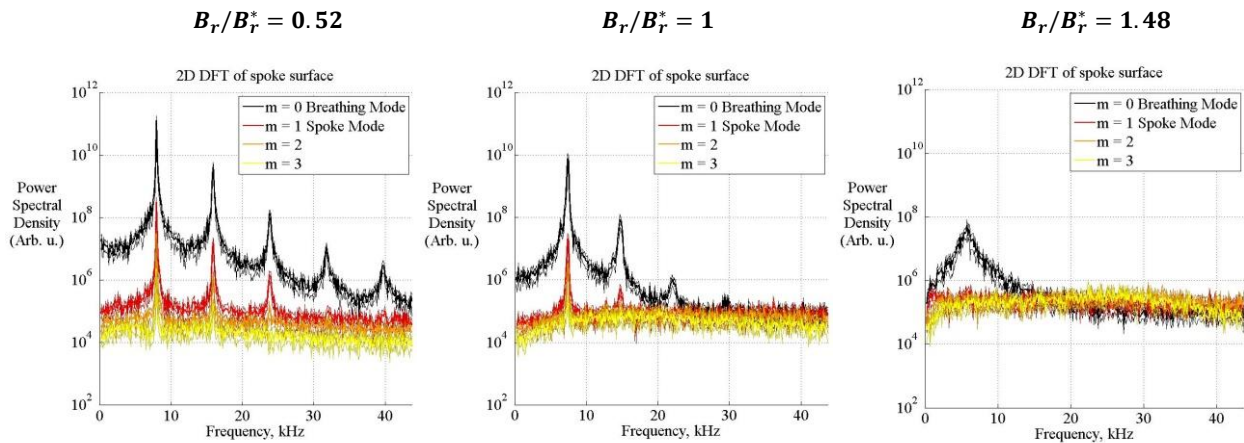


Figure 7.9. Power spectral density of FASTCAM intensity for the $m = 0$ breathing mode and spoke modes $m = 1$ through $m = 3$ for the 150-V, 10-mg/s operating condition of the H6 at three magnetic field settings. The thruster exhibits an extremely strong $m = 0$ mode at $B_r/B_r^* = 0.52$ (left). Some azimuthal perturbations visually appear at the nominal condition $B_r/B_r^* = 1$ (middle). Azimuthal perturbations are stronger but still dominated by the global oscillation at $B_r/B_r^* = 1.48$ (right).

A little more information can be gleaned from the "spoke surface" in the time domain at all three field settings in Figure 7.10. It is a density plot where the x axis is time and the y axis is the azimuthal position on the thruster (clock position). A vertical strip represents the light intensity across all azimuthal positions at a fixed time, and a horizontal strip represents the light intensity at a fixed location as a function of time. The thruster is clearly in a very coherent and high amplitude global oscillation at the low field setting. Global mode still dominates at the nominal setting, but it is noticeably weaker and less coherent. By the high-field setting, the thruster has apparently transitioned to another regime with relatively strong but incoherent azimuthal perturbations. It is distinct from the local mode described by Sekerak since there are no coherent, traveling spoke structures, which would be visible at diagonal areas of color in the plot. This provides a little more evidence than the power spectra alone that the transition observed at this operating condition may be in some sense a different type of transition than what was observed at the higher voltage and flow rate conditions.

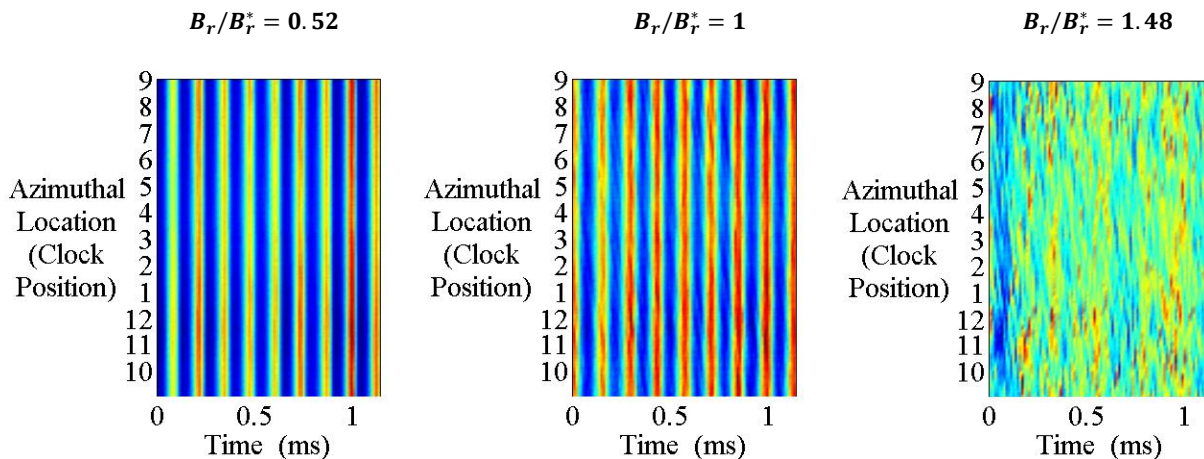


Figure 7.10. The spoke surface at low-field condition (left) exhibits a strong global mode oscillation. Azimuthal perturbations begin to appear by the nominal setting (center) and intensify at the high-field setting (right).

The power spectrum of the discharge current under 50 KHz in is similar to the $m = 0$ mode of the FASTCAM spectra. This is expected in light of previous work showing that the two signals are strongly correlated and one can even be used as a proxy for the other (see the background Subsections 2.6.1 and 2.5). There are, however, subtle differences such as the different noise floor and larger number of harmonics in the FASTCAM spectrum for the low-field condition (five compared with four in the discharge current spectrum in Figure 7.11). The cause of the differences is not clear, but the simplest explanation may simply be that the data do

not correspond to exactly the same time, and the thruster's operation may have drifted slightly between the two. Furthermore, they were analyzed with separate routines using different parameters for time-series length, windowing, and averaging over many spectra. The FASTCAM recorded about 0.19 s of data (filling the memory of the camera), while about for 1 s of the 60-s discharge current dataset was analyzed.

With a sample speed of 30 MHz instead of 87.5 kHz, the discharge current spectrum of Figure 7.11 also contains information about higher frequency oscillations not contained in the FASTCAM spectra. There are small peaks near 100 kHz that intensify slightly as the magnetic field rises. The source of these peaks is unknown. There are also some features above the noise near 1 MHz that are roughly invariant to the magnetic field setting. These oscillations in discharge current could be related to the high frequency oscillations summarized by Choueiri in a review paper on Hall thruster oscillations [4], though the oscillations described are predicted and observed to depend on magnetic field magnitude.

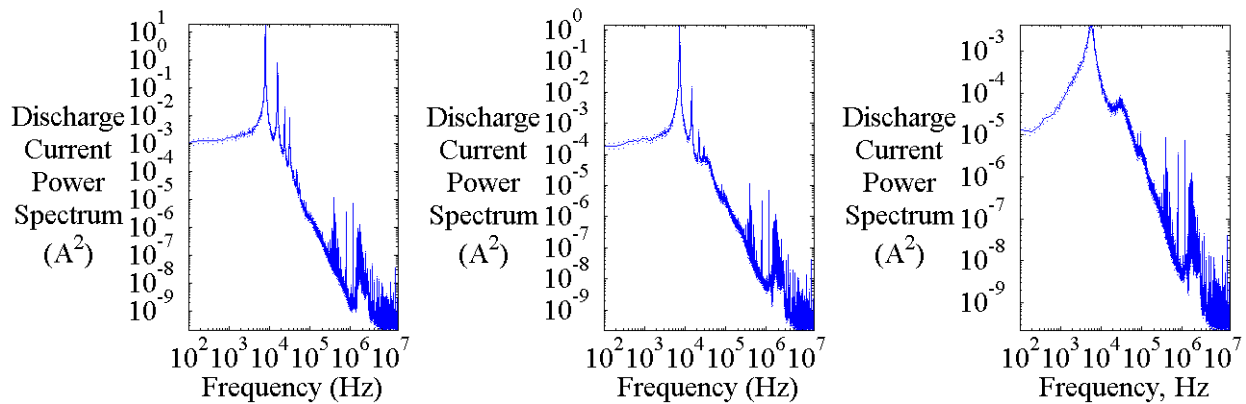


Figure 7.11. Discharge current power spectra for the low-field setting with $B_r/B_r^* = 0.52$ (left), the nominal $B_r/B_r^* = 1$ setting (center), and the high-field setting with $B_r/B_r^* = 1.48$ (right). The current spectra are similar to the $m = 0$ mode but also captures higher-frequency features.

Figure 7.12 shows that the discharge current oscillation amplitude changed substantially during the magnetic field sweep. Similar to the magnetic field sweep curves reported by Sekerak, oscillation amplitude changed by a factor of 10 between the high and low-field conditions, with the highest oscillation amplitude at the low-field setting. Unlike Sekerak's and Brown's results, the mean discharge current changed little for the magnetic field sweep at the 150-V, 10-mg/s operating condition. Mean discharge current actually increases with magnetic field magnitude from about 9 A at the low-field setting to about 9.5 A at the high-field setting.

These results for oscillation amplitude and mean current significantly differ from both Sekerak's and Brown's findings. They found two distinct modes where high mean discharge current was coupled with high oscillation amplitude and low mean current was coupled with low oscillation amplitude. The magnetic field sweep of this campaign showed a change in oscillation amplitude and coherency similar to the previous results, but the behavior of the mean discharge current is completely different. The mean discharge current stays approximately constant and high oscillation amplitude is actually coupled with slightly lower mean current. It is interesting to note that describing the oscillation amplitude in terms of +/- RMS is somewhat misleading since the maximum of the oscillation tends to be further from the mean than the minimum.

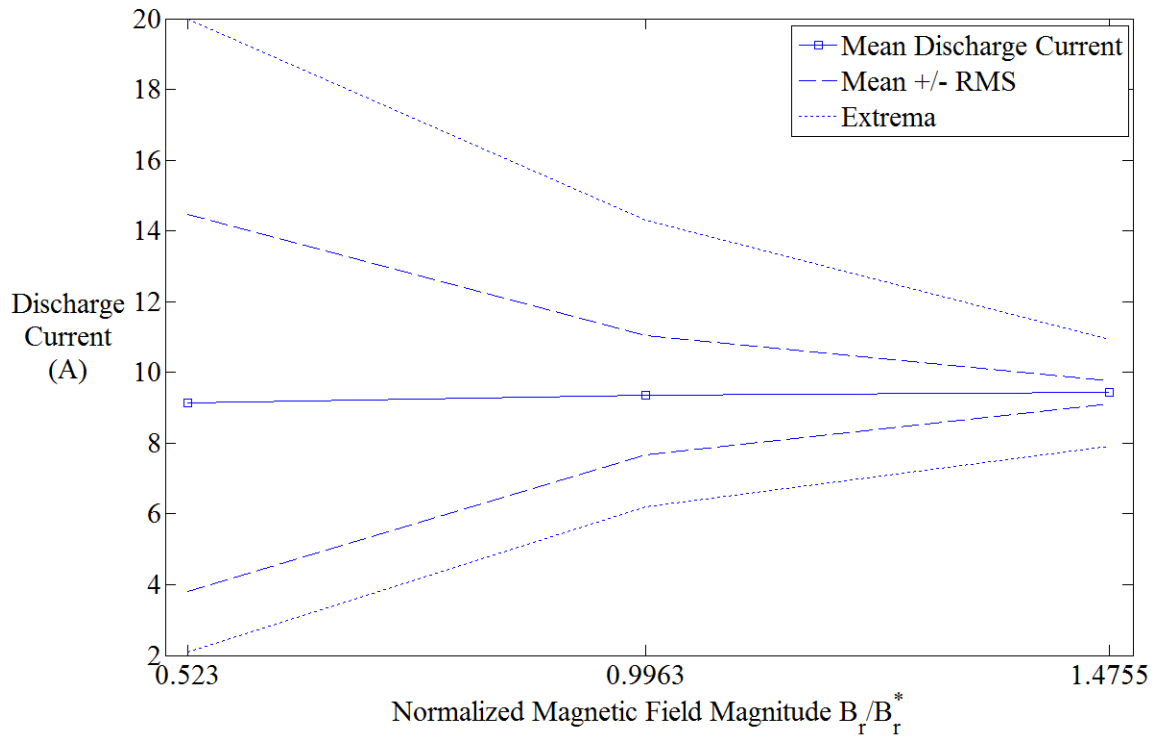


Figure 7.12. The mean discharge current (squares connected with solid line) is approximately constant while the oscillation amplitude, measured by RMS (dashed lines) or extrema (dotted lines), changes drastically during the magnetic field sweep.

Comparing this magnetic field sweep with Sekerak's criteria for H6 mode transitions introduced in Subsubsection 2.6.1.2, the oscillation amplitude has undergone a transition between the thresholds for the two modes. The mean discharge current satisfies the criterion for local mode at the same magnetic field setting where the oscillation amplitude satisfies the criterion for global mode and vice versa. The criterion for the FASTCAM peaks is tricky to

assess because the $m > 0$ peaks at low and nominal field appear to only come from smearing from the $m = 0$ peaks, while the $m > 0$ "peaks" at the high-field condition are more like white noise. With that in mind, the spectra roughly imply that the global mode condition is satisfied at low and nominal settings, and the behavior satisfies neither condition at the high-field setting (i.e. it would be considered in the transition region). In addition, the oscillations change from relatively coherent to chaotic.

The evidence seems to imply that the thruster has undergone a transition, but possibly a different kind of transition than described by Sekerak and Brown with local/global mode and low current/high current mode terminology. Considering the parallels in changes to the oscillation spectra, amplitude, and mean discharge current, the transition here could be related to the transition observed by Gascon et al. between the regimes 4 and 5 described in reference [53]. That transition occurred during a sweep of discharge voltage at higher than nominal voltage, not during a sweep of magnetic field at lower than nominal voltage (the conditions of the current experiment). It is unknown how the two observations are could be related. More investigation into the thruster operating modes at different operating condition settings is necessary to fully characterize thruster operation, let alone understanding the physics involved with the transitions.

7.2.2.2 Laser-Induced Fluorescence

The measured ion velocity distribution at the exit plane ($z = 0$ mm) for the low-field magnitude condition of $B_r/B_r^* = 0.52$ ($I_{IM} = 1.8$ A) is plotted in Figure 7.13. A Tukey window with $r = 0.5$ and a chunk size about 1 ms were used in the analysis. The oscillation is somewhat similar to that observed at the nominal field, but there are a few notable differences in the VDF. Due to the high amplitude oscillation of the mean velocity, the time-averaged distribution in this condition is by far the broadest found in this campaign, approximately spanning from 0 km/s to 12 km/s. The instantaneous velocity distribution spread averages about 4 km/s, much narrower than the time-averaged spread. An ion population first forms near 3 km/s and then steadily increases in mean velocity with a slightly decreasing spread. The highest ion density occurs near the minimum of mean velocity and gently decreases until the maximum mean velocity. After the VDF reaches a maximum in mean velocity, the population at all velocities vanishes almost within the noise floor (note the blue vertical cross sections of near-zero signal between the red pulses of high signal intensity). After a short time, the population appears again at low mean velocity and repeats the previously mentioned trends. A similar result was reported by Diallo et

al. [11] using a heterodyne TRLIF technique, lending credence to the results of both new techniques (cf. FIG. 4 in the reference). This effect may reflect a depletion of the metastable population probed by LIF, not necessarily a depletion of the ion density in this very strong breathing oscillation. More data are necessary to make a conclusion.

The ion VDF at the high-field condition $B_r/B_r^* = 1.48$ ($I_{IM} = 5.5$ A) has completely different behavior. A rectangular window and a chunk size about 1 ms were used in the analysis presented in Figure 7.14. The population density remains steady at all times and the mean velocity oscillates chaotically within only a small range of about 1 km/s between 7 km/s and 8 km/s. The VDF FWHM is nearly constant at about 2 km/s.

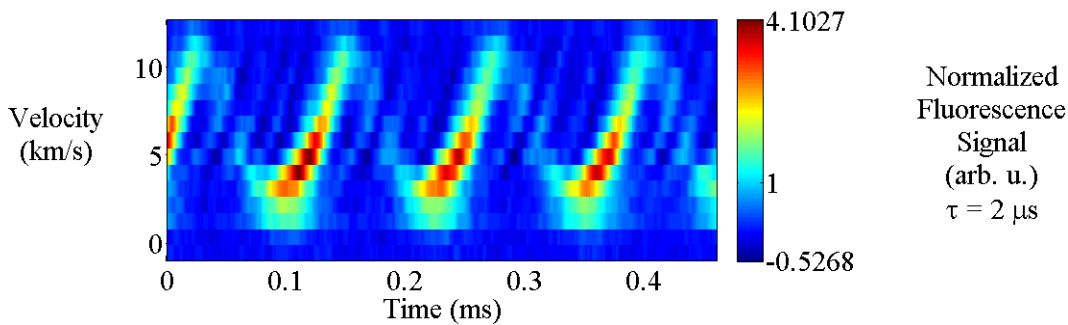


Figure 7.13. Ion velocity distribution as a function of time at $z = 0$ mm for the 150-V, 10-mg/s operating condition of the H6 with $B_r/B_r^* = 0.52$. Note that the metastable ion population probed appears to collapse almost entirely to the noise floor after the distribution reaches a maximum in mean velocity (visible as vertical stripes of mostly dark blue color).

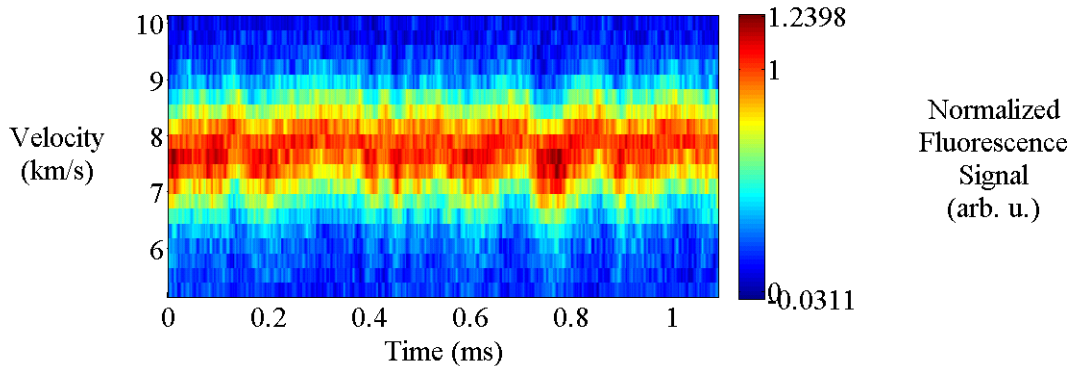


Figure 7.14. Ion velocity distribution as a function of time at $z = 0$ mm for the 150-V, 10-mg/s operating condition of the H6 with $B_r/B_r^* = 1.48$. The mean velocity oscillates nonperiodically and within a smaller range than it does at the lower field conditions.

The stark change in ion dynamics clearly indicates that different processes are occurring at the different magnetic field settings. More complete spatial maps of time-resolved LIF

captured on other Hall thrusters with stabilized oscillations have shown an oscillating electric field and ionization front apparently caused by breathing mode oscillations [80], [81]. The H6 campaign did not include sufficient spatial points to make the same conclusion, but it is reasonable to speculate that the low-field oscillations of the H6 are similar, given the similarity between the single spatial point VDFs reported in the references and the low-field and nominal-field VDFs presented here.

Apparently, breathing oscillations are stabilized with higher field conditions and different processes begin to dominate. Sekerak hypothesizes that the stabilization of the breathing mode gives rise to the conditions necessary for spoke propagation since spokes appear as the breathing mode amplitude decays [9]. The results of this campaign imply that stabilization of the breathing mode alone is not a sufficient condition for spoke propagation since coherent spokes were not observed in this campaign even with stabilized breathing oscillations.

7.2.3 Validation of TFLIF

Validating the H6 results is important because TFLIF is a new technique never before applied to Hall thrusters. The TFLIF system was first validated with a hollow cathode test bed using two types of controlled discharge current oscillations: periodic oscillations and a sinusoidal oscillation with randomly varying period (the subject of Chapter 6). The validation arguments presented here involve a similar series of benchmark tests to confirm that TFLIF results agree with other measurements in cases where the agreement is theoretically expected. No benchmark or series of benchmarks provides a logically deductive argument that results are always flawless, but together they provide strong evidence that the results are credible.

7.2.3.1 Average LIF profile comparison

The first test is to compare the time-averaged ion VDF from TFLIF with the traditional time-averaged ion VDF measured with a lock-in amplifier. An example of this comparison is shown for the high-field case ($B_r/B_r^* = 1.48$) in Figure 7.15. The average profile from TFLIF agrees with the lock-in amplifier profile within the noise of the lock-in amplifier measurement, which is roughly apparent because there are many closely spaced points with almost zero signal near the edge of the profile. In addition, the TFLIF profile is virtually identical to the triggered average LIF profile, with a mean absolute residual (the average of the absolute value of the

difference between the two TRLIF measurements) of 0.16% of the peak value. These results are good evidence that TFLIF is not introducing systematic error, at least in the average LIF profile.

This example is only for the high-field condition, but all other points presented have similar results. The worst agreement was found at the low-field condition, shown in Figure 7.16. The profiles do qualitatively agree, but it is not clear that the difference is within the noise of the lock-in amplifier. The issue stems primarily from laser tuning issues while acquiring that point and, to a lesser degree, the poor velocity resolution in the lock-in amplifier profile in this case. The laser had problems operating in a single mode at that time, and therefore the LIF profiles may not be entirely accurate due to mode completion and power fluctuations. The qualitative shape shared by all profiles is likely approximately accurate nonetheless.

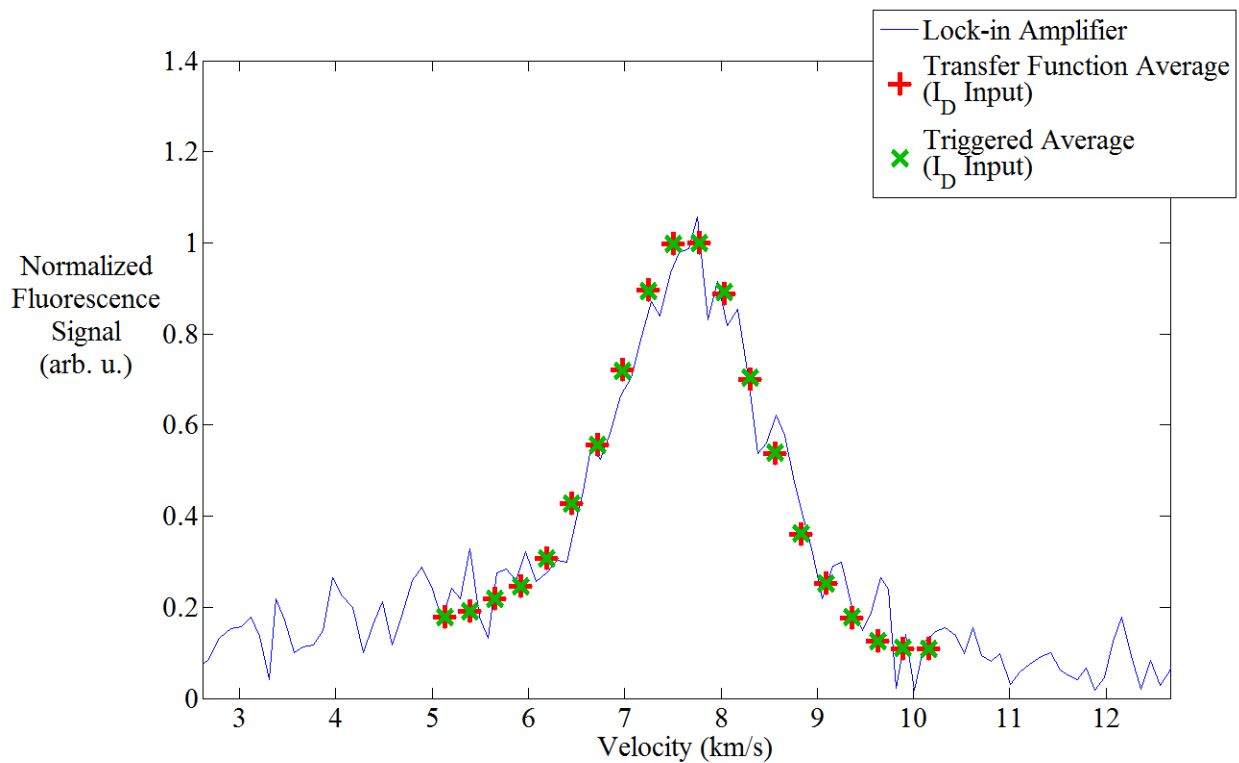


Figure 7.15. A comparison showing the time-averaged ion VDF from TFLIF agrees with a traditional lock-in amplifier measurement of the time-averaged VDF. This example is for the $z = 0$ mm, $B_r/B_r^* = 1.48$ data set.

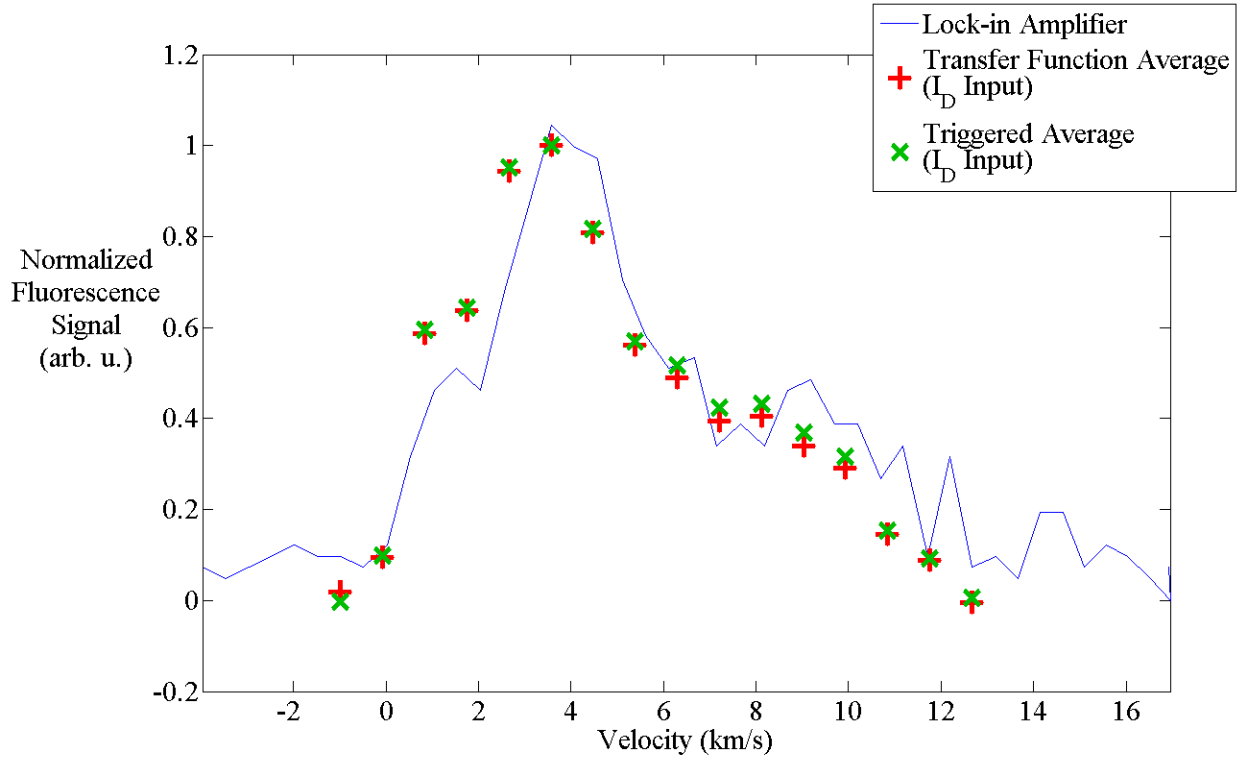


Figure 7.16. A comparison of time-averaged LIF profiles for the $z = 0$ mm, $B_r/B_r^* = 0.52$ dataset. This dataset gave the worst agreement of the points in this paper, but the time-averaged ion VDF from TFLIF still generally agrees with a traditional lock-in amplifier measurement of the time-averaged VDF. The relatively poorer agreement is likely due to trouble with laser tuning while acquiring this point.

7.2.3.2 Direct Comparison of TFLIF with the Triggered Average Signal

Even if the time-averaged profile from TFLIF is accurate, the time-resolved features could be distorted, so further benchmarks are necessary. A second test that is useful for periodic oscillations is to compare the time-resolved TFLIF signal with the triggered average LIF signal. As discussed in Subsection 4.3.4, the two analysis techniques will theoretically converge to the same result if the system is linear and the oscillation is periodic. They converge because in that case the assumptions of both techniques are satisfied and they converge to the exact TRLIF signal. This is a useful consistency check because if the two signals do not converge for periodic oscillations (conditions where the triggered average is guaranteed to work by Subsection 4.3.4), then the transfer function average has failed. This benchmark was demonstrated with a sinusoidal hollow cathode oscillation in Section 6.3, but in general is not possible with all unperturbed Hall thruster operating conditions since oscillations are not always periodic.

The direct comparison benchmark is possible for the low-field setting at the 150-V, 10-mg/s operating condition since the oscillations are nearly periodic. The comparison, demonstrating that the two signals are nearly identical, is shown in Figure 7.17. This demonstrates that the two signals from the Hall thruster do approximately converge as expected, and therefore the results are consistent with the hypothesis that the transfer function average works in this case.

The results do not converge for one particular feature. Low amplitude, high frequency oscillations appear in the transfer function density plot as thin, teal stripes that are approximately parallel with the higher amplitude bursts of LIF signal. These stripes are absent in the triggered average. Both signals have approximately the same average signal level at times between the main pulses, but the triggered average has a more uniform signal level, while the transfer function average oscillates above and below the average level. More testing with a larger dataset (to better remove noise) and more analysis parameters is required to determine whether the detected oscillation is an artifact of the signal processing or a physical oscillation that is incoherent with the trigger and therefore not captured by the triggered average.

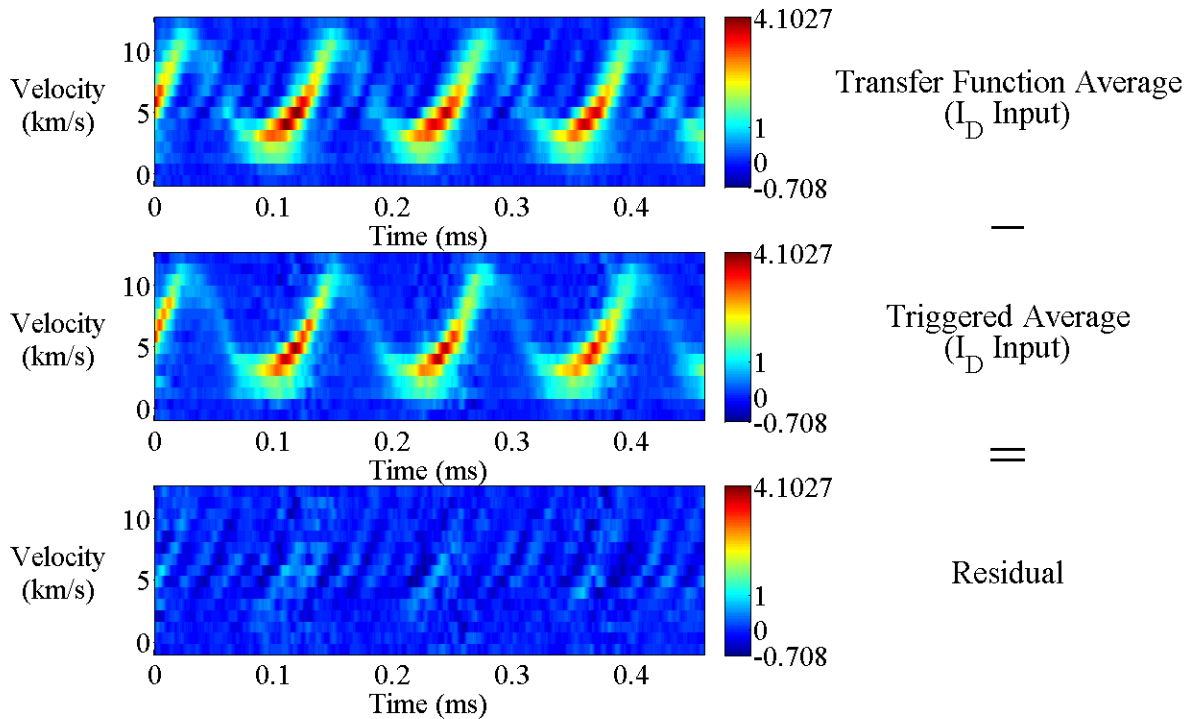


Figure 7.17. The TFLIF and triggered average signals are nearly identical at the low-field setting, as expected if the assumptions of both averaging techniques are satisfied.

The triggered average does not converge to the exact TRLIF signal in the case of nonperiodic oscillations. It typically exhibits an unphysical decay in the oscillation amplitude as the many time-series in the average are in phase at $t = 0$ but drift further out of phase from each other as time goes on. The oscillation is so nonperiodic for the high-field setting $B_r/B_r^* = 1.48$ data set that the triggered average barely detects any oscillation, as shown in Figure 7.18, and there are clear systematic differences with the TFLIF signal, as shown by the patches of blue and teal in the residual. Comparing the TFLIF signal with the triggered average signal does not provide a useful benchmark since there is no expectation of a relationship between the two signals. Hence there is a need for other benchmarks, especially ones that work with nonperiodic signals. The following two subsections deal with benchmarks that have a more general applicability.

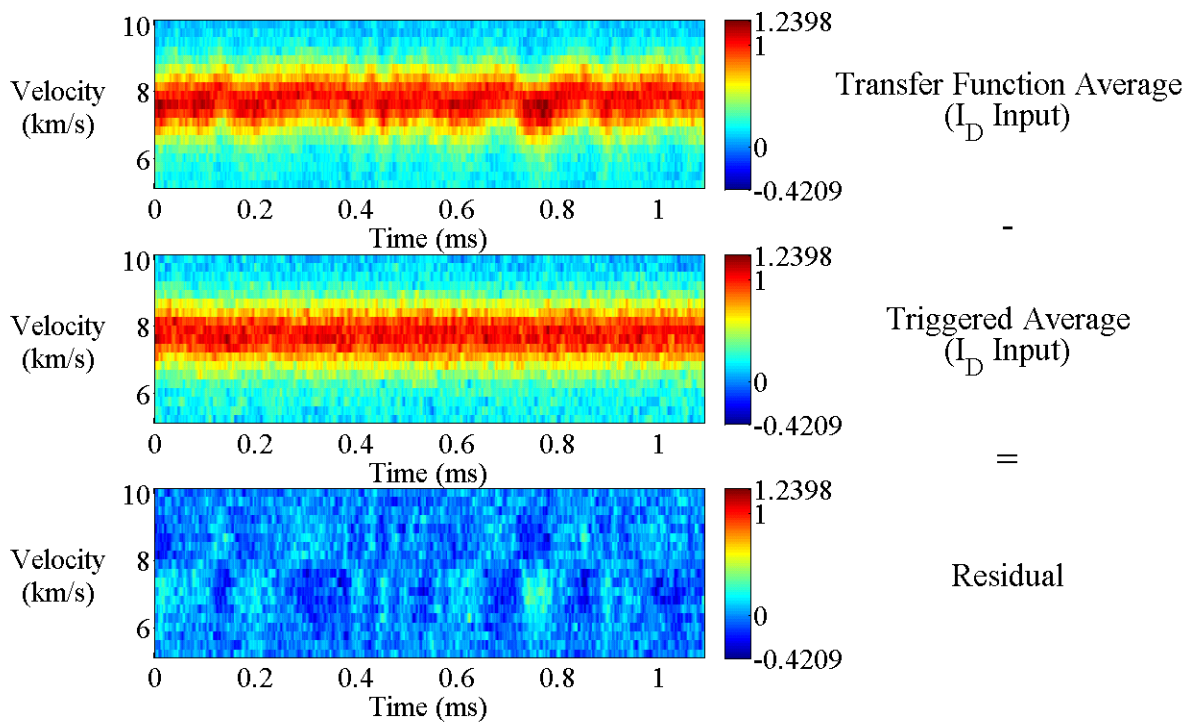


Figure 7.18. A direct comparison between the TFLIF signal and the triggered average LIF signal for the $z = 0$ mm, $B_r/B_r^* = 1.48$ dataset shows significant systematic differences. The triggered average shows a chaotic oscillation in the mean velocity but the triggered average fails to detect nonperiodic oscillations, hence another benchmark is required for nonperiodic oscillations.

7.2.3.3 General-Purpose Benchmark

An idealized analysis in Subsection 4.3.5 shows that the transfer function's characteristic LIF output signal theoretically converges to the triggered average LIF signal if the triggered average discharge current is used as the input signal to the transfer function. This fact provides a general-purpose benchmark that can be used for all signals, periodic or not. An example of this comparison is shown in Figure 7.19, which shows that the average transfer function makes an excellent reproduction of the triggered average signal with only random noise and no systematic differences in the residual. This helps confirm that the assumptions that the transfer function analysis is based on are valid (e.g. that the system is linear, that reasonable analysis parameters were used, and that the average transfer function has sufficiently converged to the exact transfer function to give accurate results). If those assumptions were invalid, then we would expect some artifacts to be present in the transfer function's reproduction of the triggered average signal, leading to systematic differences in between it and the triggered average result.

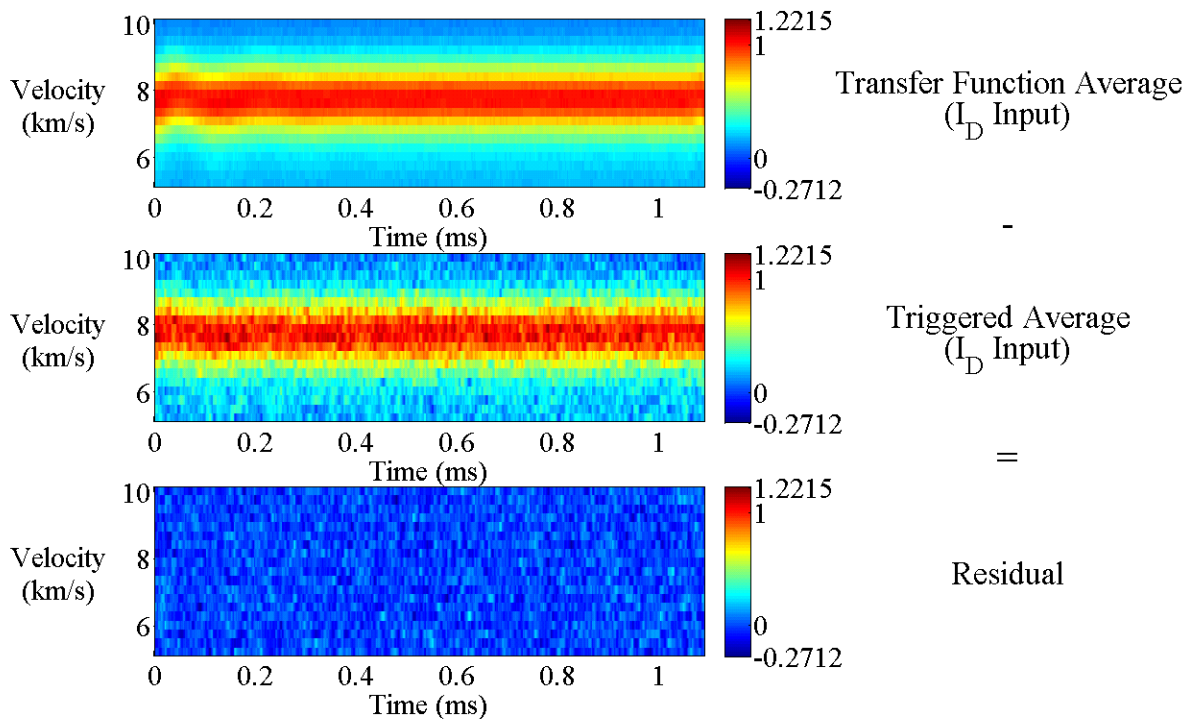


Figure 7.19. An example of the general-purpose benchmark with the $z = 0$ mm, $B_r/B_r^* = 1.48$ data set. The transfer function faithfully reproduces the triggered average LIF signal when the triggered average discharge current is used as input to the transfer function.

7.2.3.4 *Different Input Signals*

A final benchmark is to compare the resulting characteristic TFLIF signals generated from two different input signals. For each wavelength, an average empirical transfer function has been generated using the same 60-s dataset containing the TRLIF signal in the PMT signal and a system input of either the discharge current or the simultaneously measured floating probe potential. There is not necessarily any one "special" input signal for which the transfer function exists; any signal measured from the thruster could potentially have a linear relationship with the TRLIF signal, and therefore contain the information necessary to map the input signal to the output signal via a transfer function. The transfer function for one input signal will in general be different than the transfer function for another input signal, but two simultaneously measured input signals will both map to the same output signal if both input signals have a linear relationship with the output signal.

That is the basis for this benchmark. If the characteristic TFLIF signal from the two input signals are indistinguishable, then it is consistent with the idea that both input signals likely have a linear relationship with the output fluorescence signal and therefore the signals are an accurate measurement. If they fail to recover the same output signal, however, at least one or both of the input signals may not actually have a linear relationship with the output signal, but there is not necessarily evidence to conclude which one is accurate (if either one).

Figure 7.20 compares the TFLIF signal for the low-field setting ($B_r/B_r^* = 0.52$) and analysis with a chunk size of about 1 ms with a Tukey window. Both analyses capture the main features discussed in Subsubsection 7.2.2.2 and the residual is mostly near zero, verifying that both input signals do converge to essentially the same result.

The TFLIF signal from the discharge current input more distinctly captures the relatively high-frequency oscillation shown as a series of light blue stripes that are parallel with the more intense bursts of signal. Though they both capture it to some degree, the stripes are more diffuse with the floating probe input, which gives rise to the stripes in the residual instead of the usual random noise pattern.

We can only speculate on possible causes for the discrepancy based on this evidence. The linear relationship between input/output may not be perfect for (at least) one of the inputs, leading to a distorted signal. The floating probe input seems likelier to be the problem since the signal is muddier, but we cannot make conclusions with the available data. Another possibility is

simply that the TFLIF signal from the floating probe input may be distorted because the input signal is comparatively noisier than the discharge current. The noisier input could lead to a noisier average transfer function or directly contribute to distortion of the TFLIF signal from the noise in the input signal used to generate the characteristic TFLIF signal.

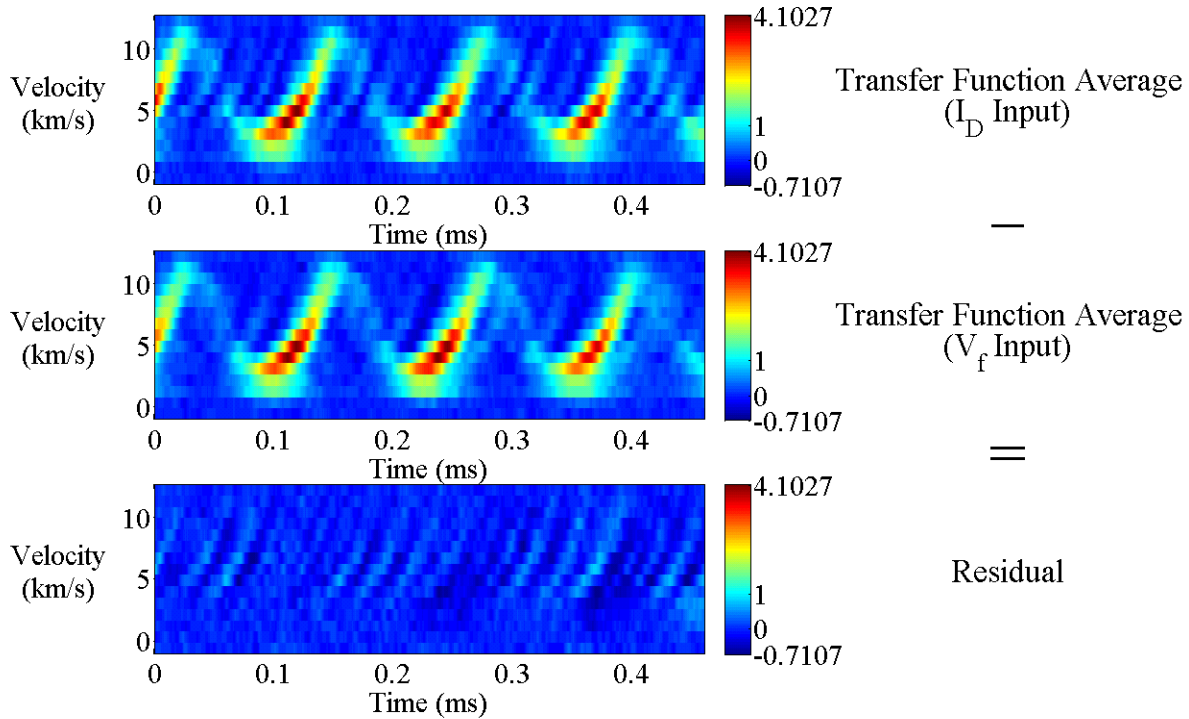


Figure 7.20. A direct comparison of TFLIF signals from discharge current (top) and floating probe voltage (center) input signals for the low-field setting $B_r/B_r^* = 0.52$ shows small systematic differences in the residual (bottom).

Figure 7.21 compares the TFLIF signals of both inputs for the high-field setting with ($B_r/B_r^* = 1.48$). Analysis used a chunk size of about 1 ms with a Tukey window. The figure shows that both analyses capture the general character of the chaotic oscillation. They also capture the same oscillations in mean velocity and relative metastable ion density, especially after the first 0.1 ms of the window. This is somewhat apparent in the VDF heat maps, but is particularly visible in the plots of mean velocity and relative metastable ion density in Figure 7.22, calculated from the data of Figure 7.21. The mean velocity is calculated via the first moment of the distribution at all time steps. The relative metastable ion density is calculated by numerically integrating over the distribution profile using the trapezoid rule. The two output signals agree even with two completely different input signals (see Figure 7.23).

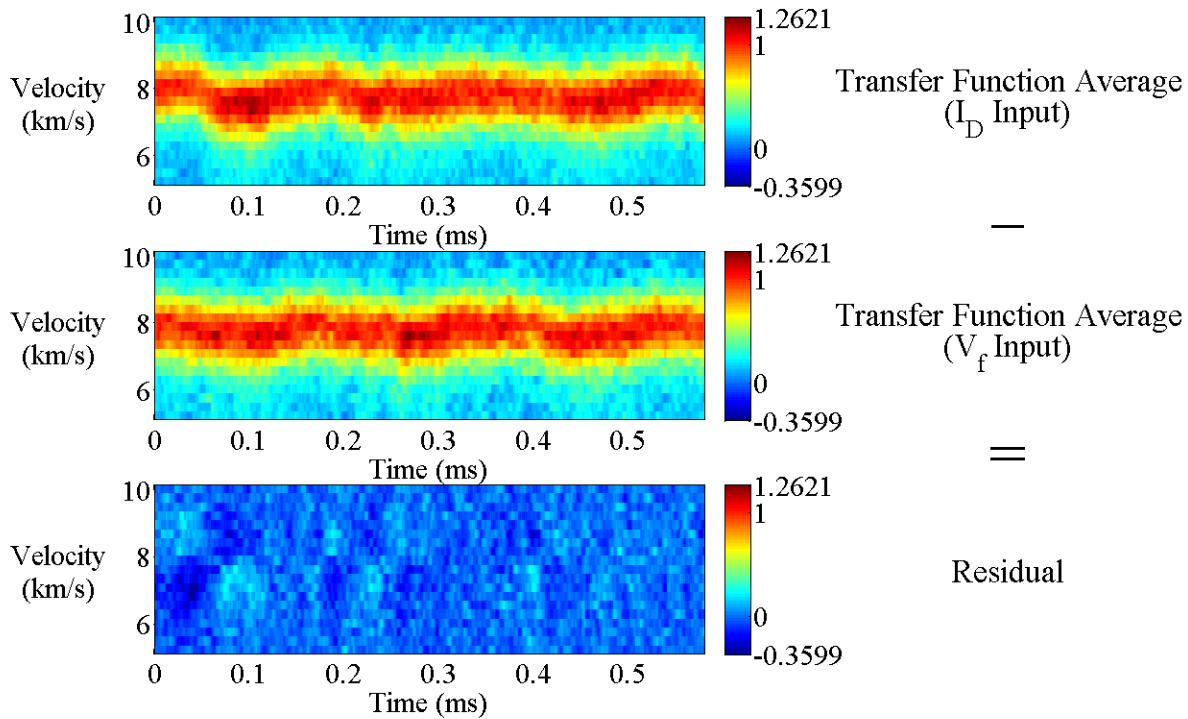


Figure 7.21. The TFLIF signals from the discharge current (top) and the floating probe voltage (center) input signals capture the same general features for the high-field setting $B_r/B_r^* = 1.48$. The residual (bottom) is almost entirely random noise in the second half of the window, while the first half of the window shows differences in the amplitude of the mean velocity oscillation.

The oscillations of mean velocity and metastable ion density in Figure 7.22 do not agree in about the first 0.1 ms of the plotted window, but they do agree reasonably well otherwise. Further investigation is needed to fully characterize how these minor errors arise, but, as with the minor differences at the low-field case, we speculate on some possible culprits. The differences may indicate small nonlinearities in the system relating one or both of the inputs to the output. In particular, the floating probe voltage is a local measurement while the discharge current is the net current passing through the thruster globally. It is possible that the probe could retain a linear relationship with the output signal necessary to capture azimuthal waves even if the linear relationship breaks down for the discharge current input. In fact, one of the reasons that the floating probe was captured in this campaign was to have an alternate input signal sensitive to azimuthal oscillations in case the linear relationship breaks down for the discharge current input with azimuthal spokes. Another possible cause of the discrepancy may be artifacts resulting from a suboptimal choice of chunk length, windowing, or other analysis parameters. Another

possibility is that the signals are noisy enough that the average transfer functions have not sufficiently converged to the exact transfer function in order to give fully accurate output signals.

The comparison of TFLIF signals with different input signals does generally confirm the expected relationship between the two output TFLIF signals. Though there may be questions about the accuracy of some particular features, the fact that both input signals generate equivalent output signals provides evidence validating the results.

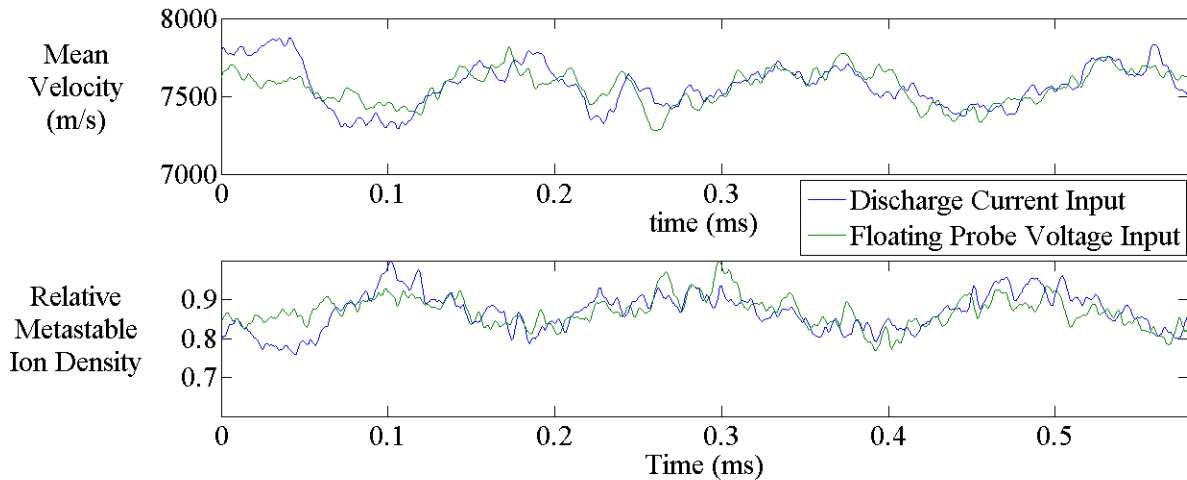


Figure 7.22. The mean velocity (top) and the relative metastable ion density (bottom) of the TFLIF signal from the discharge current (blue) and floating probe (green) inputs agree for most of the window except the first 0.1 ms. These VDF properties were calculated from the dataset of Figure 7.21.

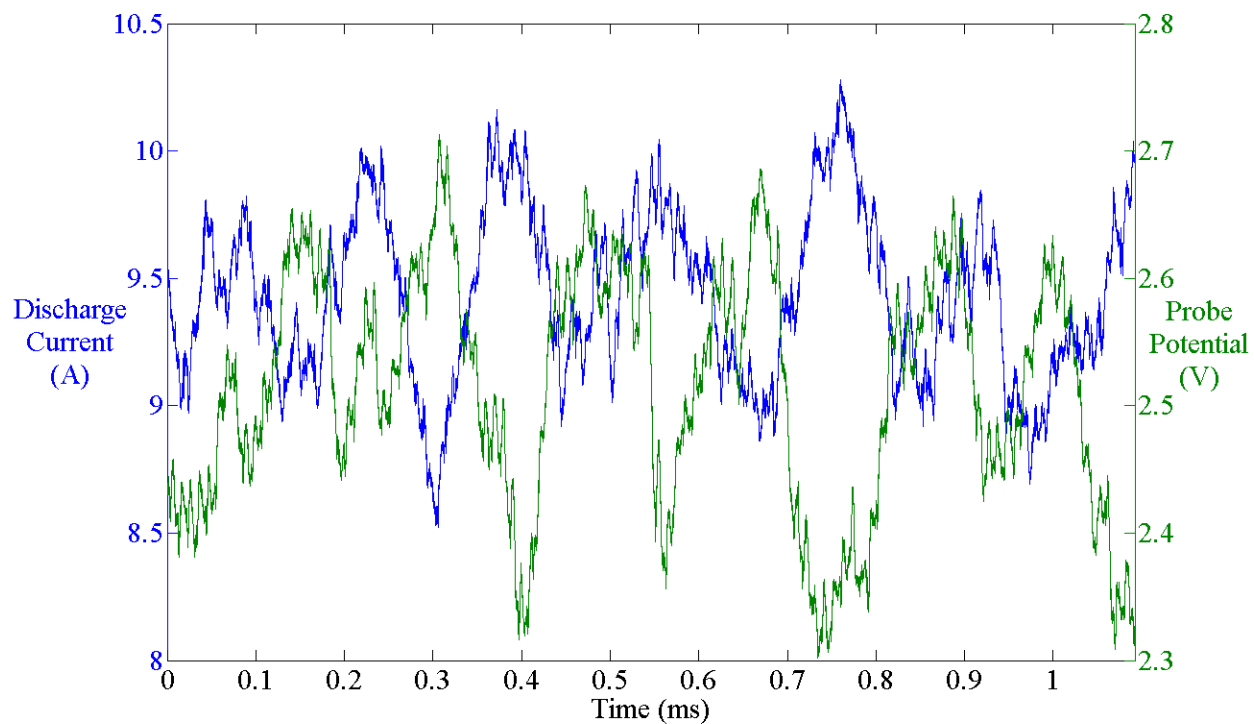


Figure 7.23. The input signals used to generate the TFLIF signals in Figure 7.21 are completely different, but the TFLIF signals output from their respective transfer functions are very similar.

Chapter 8

Discussion and Conclusion

“Every new beginning comes from some other beginning's end.”

– Lucius Annaeus Seneca³

Having described the theory of the TFLIF technique (Chapter 4) and its implementation (Chapter 5), provided arguments validating results (Chapter 6), and an initial campaign demonstrating the technique with a Hall thruster (Chapter 7), we enter the final chapter of this dissertation. First, Section 8.1 evaluates the TFLIF technique against the many other techniques that have been applied to TRLIF experiments on Hall thrusters in the past few years. Each technique has a unique set of advantages and disadvantages that may make it attractive for a certain niche of applications. The benchmarks used to help validate TFLIF signal are listed and the conclusions of validation tests are discussed in Section 8.2. The key conclusions regarding Hall thruster physics that can be made from the results of this dissertation are discussed in Section 8.3. Section 8.4 lists several possible branches of future research that are almost immediately actionable and likely to give fruitful results. Finally, Section 8.5 briefly summarizes how this dissertation addressed the problem and aim that were stated in the introduction.

8.1 Comparison of Features and Specifications of Time-Resolved LIF Techniques

The TRLIF techniques discussed in Section 3.3 all share a common limitation to periodic or at least quasiperiodic oscillations, at most allowing for some variation in the period of oscillations but implicitly assuming a constant amplitude and waveform shape. This common limitation was used to argue in that background section for the need for a TRLIF technique capable of measuring the TRLIF signal for a more general class of oscillations. Developing and demonstrating such a new technique was the aim of this dissertation.

³ Commonly attributed to Seneca. Also a prominent line in the song “Closing Time” by Semisonic, which is suggested as an appropriate accompaniment to this chapter. I found it cathartic while finishing writing, at least.

The capability to recover the signal for a more general class of oscillations is only one of many important features to consider when evaluating techniques. This dissertation would be incomplete without discussing the trade-offs involved in achieving that feature. Many highly talented scientists and engineers have worked on developing each one of the techniques, and therefore it is no surprise that each technique has a distinct set of advantages and disadvantages and no one technique dominates in all of the important features. A summary of the key features and specifications of the four main techniques used for TRLIF in Hall thrusters and related EP devices is given in Table 8.1. The rationale used to generate the table is discussed in further detail in the following subsections, but the basic conclusion about how the techniques generally compare is discussed here.

Table 8.1. Summary of the five most relevant features of four TRLIF techniques used with Hall thrusters or similar EP devices. The most attractive features of each category are highlighted in green. Tilde denotes an estimate that is likely accurate within a factor of about 3 or better.

	Synchronized Photon Counting	Sample-and-Hold	Heterodyne	TFLIF
Acquisition Time Per Wavelength	~5 m	~10 s	~1 s	1 m
Signal Bandwidth (kHz)	~1000	~100	~24+	65+
Post-Processing	minimal	minimal	minimal	extensive
Monetary Cost	~\$10,000	~\$10,000	~\$1,000	~\$10,000
Assumption of the Averaging Scheme	Periodic Oscillations	Quasiperiodic Oscillations	Periodic Oscillations	Linear map from input signal to output signal

It could be argued that each technique is best suited for a particular niche of applications. The high bandwidth makes synchronized photon counting the best choice for high frequency, periodic oscillations. With a high resolution in velocity space implied by the short acquisition time, the sample-and-hold technique may be best to observe finer details of the LIF profile undergoing quasiperiodic oscillations with minor drift or variation. The heterodyne technique offers a similarly high resolution in velocity space and can be performed at low cost, but is likely less tolerant of variation in the oscillations. The TFLIF technique has relatively mid-range

specifications in every category and could be reasonable choice for general-purpose use, but it is only extremely attractive for use with nonperiodic oscillations.

8.1.1 Acquisition Time per Laser Wavelength

The acquisition time per laser wavelength that is required for a measurement varies by about a factor of 100 depending on the TRLIF technique. Note that acquisition times are not quite directly comparable since differences in the plasma source and optics may influence the necessary acquisition time. Nonetheless, the wide range in the differences in acquisition time is likely due largely to the measurement techniques themselves.

Synchronized photon counting is by far the slowest due to the small time bins over which averaging is performed, the lack of analog signal processing to reject part of the noise spectral density before a main averaging technique, and the requirement of two triggers per complete averaging cycle (see Appendix B). The sample-and-hold and heterodyne techniques are capable of fast signal recovery in nearly the same time as a time averaged experiment. Both schemes use commercial lock-in amplifiers to recover the signal similarly to a traditional time-averaged experiment. Unlike a traditional LIF experiment, they use multiple lock-in amplifiers simultaneously to recover the signal from different time gates (for the sample-and-hold technique) or to recover the heterodyne signal corresponding to Fourier series components (for the heterodyne technique). The sample-and-hold technique is somewhat slower than the heterodyne technique because multiple laser scans are necessary to capture all of the time gates, though adding many more lock-in amplifiers to the system could eliminate that need in principle.

The TFLIF system has midrange performance between the two extremes. Achieving a reasonable SNR requires an acquisition time of about 1 minute despite the benefits from signal conditioning of filtering and phase-sensitive detection, possibly due to a propagation of uncertainty in calculating transfer function estimators.

The acquisition time per laser wavelength directly influences the number of laser wavelengths that are practical to acquire, and therefore the resolution in velocity space of the resulting LIF profile. Hence, synchronized photon counting is typically done with about ten wavelengths and TFLIF with twenty. The sample-and-hold and heterodyne profiles typically have 100 or more points in velocity.

8.1.2 Signal Bandwidth

The bandwidths also can only be compared approximately due to the major differences in how the data are collected and analyzed. The raw data of TFLIF are digitized at a sampling frequency of 30 MHz but TFLIF has a much lower effective signal bandwidth due to the low-pass filtering from phase-sensitive detection. The highest signal bandwidth possible for the experiments in this dissertation was estimated to be about 65 kHz in Subsubsection 4.3.2.4 using the requirement that all frequencies within the signal bandwidth be subject to a gain distortion of 5% or less. Bear in mind that this was estimated for a 2-MHz modulation frequency, which was the optimal modulation frequency with the H6 campaign. That campaign had a degraded SNR at least partially attributable to an inadvertent mismatch of the numerical aperture of the collection optical fiber and feedthrough. An experiment correcting this and (possibly) other problems with SNR may improve SNR significantly, allow a higher modulation frequency, and therefore ultimately a higher bandwidth.

Synchronized photon counting averages photon counts over bins of size 100 ns (as implemented). The distortion of a waveform is not well defined under these circumstances and will vary depending on the particular features of the recovered waveform, but it is reasonable to require at least 10 points per period with each point averaging over at most 1/10 of the period; hence we estimate a bandwidth of 1 MHz. Similarly, the sample-and-hold technique averages over a gate width of 1 μ s, leading to a 100 kHz bandwidth. The heterodyne technique is peculiar in that it does not directly sample in the time domain at all but recovers Fourier series components with the lock-in amplifiers. The most recent publication [11] was done with a 12-kHz oscillation with Fourier series components recoverable up to second order, or 24 kHz. The bandwidth of that particular measurement was then 24 kHz, but that is not necessarily the limit of the system, which could presumably also recover the second order components of much faster oscillations.

With three different methods used to estimate the bandwidths of the four techniques, there is considerable room for debate for the values of bandwidth and how they compare. In any case, the synchronized photon counting technique is the clear leader in terms of signal bandwidth or time resolution.

8.1.3 Post-Processing

One major difference between TFLIF and the other techniques that should be reiterated here is the need for extensive post-processing since raw data from the photomultiplier are digitized and saved. This implies the need for a large capacity RAID and powerful computer to perform hours of calculations for a signal time-resolved VDF profile measurement (see the considerations of Chapter 5). The other techniques record only the resulting TRLIF signal after the instruments perform the signal recovery. On the other hand, the burden of signal recovery done in post-processing enables the raw data to be analyzed with different parameters (*e.g.* Subsection 5.4.4) or different averaging techniques (*e.g.* the transfer function average versus the triggered average). The extra flexibility can be useful to set optimal analysis parameters and can help validate the resulting TFLIF signal. The need for post-processing, then, should be considered to be a trade-off rather than strictly an advantage or disadvantage.

8.1.4 Monetary Cost

The exact monetary cost of implementing the various systems will vary widely depending on what equipment is already available in the laboratory and the price of the particular instruments purchased. A general and detailed estimate of the costs to implement each technique is therefore not possible. All of the techniques except the heterodyne technique, however, require specialized equipment not normally available in a laboratory equipped for time-averaged LIF that could easily total to tens of thousands of dollars. The heterodyne technique is the most inexpensive and possibly the easiest to implement because it requires no special equipment outside of the circuit to drive periodic oscillations in the thruster.

8.2 The Development of TFLIF

The Transfer Function Laser-Induced Fluorescence (TFLIF) system was developed and tested under a variety of conditions: a hollow cathode with periodic and random discharge current oscillations as well as an initial campaign with the H6 Hall thruster. A series of benchmarks was proposed and demonstrated to help validate the signal. The benchmarks involve comparing the TFLIF signal with another signal under circumstances where the two signals are theoretically expected to converge to the same result. These cases include:

- (1) Comparing the time-averaged TFLIF profile with the traditional time-averaged LIF profile from a lock-in amplifier

- (2) For periodic signals, comparing the time-resolved TFLIF and TALIF profiles directly
- (3) For non-periodic signals, the “general-purpose” benchmark: comparing the TALIF profile with the TFLIF profile associated with the triggered average input signal
- (4) Comparing the two TFLIF profiles associated with two different input signals

We can confidently conclude that the system worked very well under the conditions of the periodic cathode oscillation and the quasiperiodic Hall thruster operating modes with low and nominal B-field settings. The benchmarks (1) and (2) demonstrated little to no systematic error (see Section 6.3 and Subsection 7.2.3). In addition to the benchmarks, the results follow several trends that can be explained physically. One trend is that the LIF signal tends to follow discharge current and frequency. Secondly, the phase of the TFLIF signal at different laser wavelengths are related in a way that makes sense physically when viewed together in a density plot or animation. For example, the Hall thruster at low B-field shows an ion VDF that oscillates between high and low mean velocities. Another example is the relationship between the two points upstream and downstream of the exit plane where the VDF spread is always about the same or greater than predicted by kinematic compression.

The results regarding the nonperiodic conditions were generally positive as well, but, until more data is collected, the conclusion must be more cautiously optimistic than for the periodic oscillations. The hollow cathode test with random oscillations showed excellent agreement with the applicable benchmarks. It also showed physically expected trends such as following the changing period of the discharge current and the amplitude varying with the period due to operating near the -3 dB point of the discharge current amplifier. For the nonperiodic Hall thruster condition, the benchmarks were generally satisfied except for some minor problems with benchmark (4) listed above. There were some small but noticeable patches of systematic differences between the TFLIF characteristic signals from discharge current and floating probe inputs. There were also small differences between their calculated mean velocities, particularly at the edge of the window. The discrepancies could be caused by a breakdown of the linear map from one (or both) of the input signals to the TFLIF signal. It seems more likely, however, that the discrepancies were caused by some artifact of the analysis that could be ameliorated with a higher SNR in the raw data, more averaging, or optimizing analysis parameters. More work is necessary to make a further conclusion.

8.3 Hall Thruster Physics

The most important result regarding Hall thruster physics is the confirmation that broad and bimodal VDFs, such as observed by Huang [45], can be caused by a relatively narrow instantaneous ion VDF that oscillates between high and low mean velocities. This implies that the thruster operating mode with high-amplitude discharge current oscillations at low magnetic field settings is likely associated with an axial oscillation of the location of the acceleration zone. In contrast, the mode with low-amplitude discharge current oscillations at high magnetic field is associated with a stabilization of the oscillation of the acceleration zone.

For now, these conclusions are well founded only for the single operating condition tested with a discharge potential of 150 V and an anode flow rate of 10 mg/s. Sekerak and Brown described two main thruster modes at a wide variety of operating conditions of the H6 (see Subsubsection 2.6.1.2). Given their results, it is reasonable to expect a similar ion VDF at the other operating conditions where these modes are observed, but a more comprehensive conclusion must wait for more results at other operating conditions. The ion VDF and its transition as a function of magnetic field may be qualitatively different at the nominal operating conditions, especially considering that the mode transition observed in this study was different from those described by Brown and Sekerak. In this study, the mean discharge current was relatively independent of magnetic field and spokes were absent. At other operating conditions, Brown and Sekerak found that high-amplitude oscillations were associated with a significantly higher mean discharge current and spokes were always present in the mode with low-amplitude oscillations of the discharge current.

The finding that the VDF spread at the point 15 mm downstream of the exit plane is about the same as predicted by kinematic compression at some times and a larger by as much as factor of about 3 at other times is a relatively minor but interesting result. It is impossible to make conclusions with the sparse data available, but we can speculate on the origin of this effect. It could be related to an oscillation in the ionization zone. The width at the downstream point was predicted based on the width at a point upstream and assuming kinematic compression is the only effect influencing the VDF between those points. If the location of ionization zone were oscillating such that ionization sometimes occurs between the two points and at other times ionization takes place completely upstream of $z = -4$ mm, that effect could explain the result.

The fact that the instantaneous VDF width stays approximately constant while the mean velocity oscillates violently supports this hypothesis. The spread of the VDF is largely set by the overlap of the ionization and acceleration zones; hence if the acceleration zone oscillates and spread remains constant, then the ionization zone should oscillate to keep about constant overlap with the acceleration zone. The results and speculation are also supported by more complete time-resolved LIF measurements on other Hall thrusters [79], [81].

8.4 Future Research

The discussions above and the dissertation at large suggest a number of possible avenues for research in the immediate future:

- (1) Completely explain and ameliorate the severely degraded signal compared with previous LIF measurements with the H6. Since the initial TFLIF campaign with the H6, it was discovered that the collection fiber feedthrough had a numerical aperture incorrectly matched with the fibers, effectively cutting off a significant portion of collected light. It is not yet clear whether that error alone fully accounts for the low SNR observed.
- (2) Repeat the H6 campaign with a magnetic field sweep at the nominal thruster operating conditions. These data will be more relevant and comparable to the results of Sekerak and others. The existence of more coherent spokes near the nominal operating conditions will allow the investigation of ion dynamics as spokes pass over the LIF interrogation volume.
- (3) Repeat the H6 campaign at the 150-V, 10 mg/s operating conditions with a more dense collection of points in axial position. This could enable the calculation of the electric field and ionization rate using a method such as developed by Pérez-Luna [104] or Spektor [105]. The time-resolved measurements of the acceleration and ionization zones may help answer the questions raised in Section 8.3 regarding possible motion of the acceleration and ionization zones. Furthermore, measurements of the VDF at many spatial locations at the high-field operating condition with a higher SNR may help to determine more clearly whether the random, small-amplitude oscillations in the mean velocity are physical or possibly the result of some artifact or noise.
- (4) Even more information can be gained with the combination of the TFLIF and the high-speed dual Langmuir probe (HDLP) systems. Previous measurements of the HDLP system with other Hall thrusters have indicated oscillations in both electron density and temperature [7], [36], with electron density in particular varying with a large enough

amplitude that Lobbia wrote of “discrete toroidal plasma emission” [36]. HDLP measurements together with TFLIF may be able to determine whether a severe drop in electron density and/or temperature explains the periodic loss of TFLIF signal observed at the low magnetic field operating condition of the H6. Conversely, TFLIF may be able to diagnose possible distortion of the time-resolved ion VDF from the presence of the HDLP. In general, TFLIF and HDLP signals together for exactly the same conditions will also provide further validating evidence for both techniques if they tell a coherent story.

- (5) Time-resolved measurements with more than one axis of velocity information could uncover more information about ion dynamics during the two main thruster operating modes. Due to the time required for time-resolved measurements, it may be ideal to improve the setup so that multiple velocity axes can be recorded simultaneously. Similar to a traditional experiment with choppers and lock-in amplifiers, it would require multiple AOMs with modulation frequencies distant enough to avoid interference, which could be difficult with the short time constants necessary for TFLIF and fairly tight bounds on acceptable time constants.
- (6) Improve the efficiency of the analysis code. For example, the code finding the modulation phase of the LIF signal currently runs with the analysis each time a data set is analyzed, taking about 10 minutes per wavelength per analysis while the total analysis can take about 1-2 hours. Analysis of the same data set is often performed many times with runs of triggered and transfer function averages often performed with different variations of analysis parameters. The modulation phase of the LIF signal in each data set of course remains the same, so it would be more efficient to find the phase once and save it for each subsequent analysis. This is just one improvement that could reduce the total calculation time by about 5-15%, or possibly more. This improvement may even enable more simultaneously active workers since the phase finding routine generally has a higher CPU utilization than the transfer function averaging. There are undoubtedly other ways to improve the efficiency of the code that could be uncovered. Code efficiency may not appear to be a critical branch of the research, but computational time is a main limitation of the TFLIF system since the analysis of a complete VDF of about 20 wavelengths can take more than 12 hours. Perhaps developing a more efficient version of the code could be a

good project for a master's degree student in computer science who is interested in scientific computing.

8.5 Conclusion

The problem that this dissertation addresses is the need to measure the TRLIF signal in nonperiodic modes of Hall thrusters without perturbing the operating conditions. The aim of this dissertation stated in the introduction is to investigate how the problem may be solved with a new measurement technique. To that end, a new technique has been developed and analyzed theoretically and experimentally tested in a number of different conditions. It has proven to be accurate in periodic and quasiperiodic cases, and there is also strong evidence that it is capable of recovering random and chaotic oscillations. Though some questions remain about the results obtained with the H6, many of them will be addressed in future work. The technique has proven to be a reliable and powerful method capable of competing with the other TRLIF techniques. In particular, TFLIF is the most promising technique that may allow measurements in borderline quasi-periodic and completely nonperiodic Hall thruster operating modes.

Appendices

Appendix A

Window Effects

This appendix summarizes the effects of different windows on the output TFLIF signal and remarks on the trends observed. The window commonly used in this dissertation is the Tukey window, which is piecewise defined in three regions by:

$$w(x) = \begin{cases} \frac{1}{2} \left(1 + \cos\left(\frac{2\pi}{r} [x - r/2]\right) \right), & 0 \leq x < \frac{r}{2} \\ 1, & \frac{r}{2} \leq x < 1 - \frac{r}{2} \\ \frac{1}{2} \left(1 + \cos\left(\frac{2\pi}{r} [x - 1 + r/2]\right) \right), & 1 - \frac{r}{2} \leq x \leq 1, \end{cases} \quad (\text{A.1})$$

where r is a selectable parameter and x is a fraction of the total time series (chunk) used in analysis. The Tukey window weighting is flat like the rectangular window in the center region, which is a fraction of $1 - r$ of the window. The wings, each taking a fraction of $r/2$ of the window, are cosine functions scaled to connect zero at the boundary with one at the center region. Note that $r = 0$ is equivalent to the rectangular window. Some examples of the Tukey window are shown in Figure A.1.

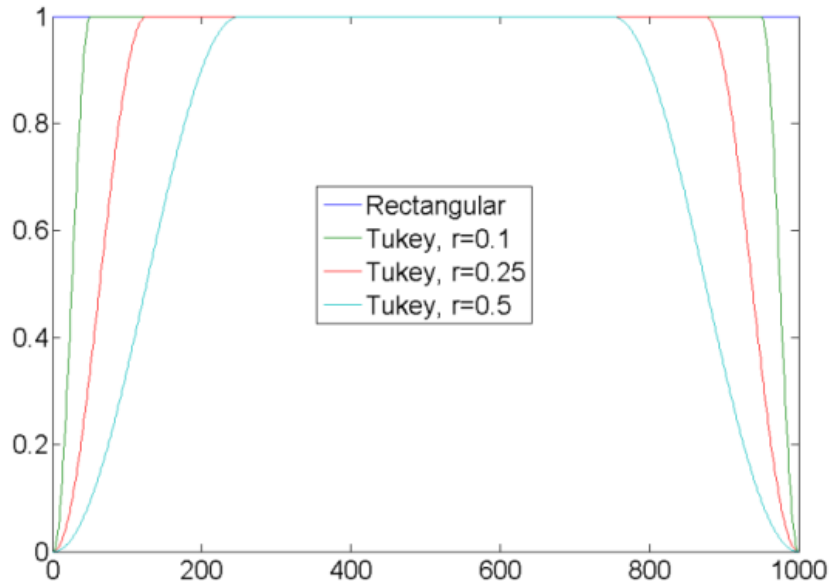


Figure A.1. An illustration of the shape of the Tukey window for a selection of r values and a time series length of 1000.

The central purpose of using a window with the discrete Fourier transform is to reduce artifacts that arise from a discontinuity at the edges, since the DFT implicitly assumes the input time series is one period of a periodically repeating signal. Hence, the weighting of all windows approaches zero at the edges and tends toward one near the center in an approximately bell-shaped curve. More subtly, the window also provides a tradeoff between a number of figures of merit such as the -3 dB bandwidth of the main lobe, the equivalent noise bandwidth, scallop loss, maximum side lobe level, and side lobe roll off. In practice, the noise floor, dynamic range, and ability to resolve closely spaced frequencies at similar or dissimilar amplitudes depends strongly on the window used.

Note that with windowing the characteristic output of Equation (4.49) becomes:

$$F_{characteristic}[k]w[k] = \mathcal{F}^{-1} \left(\mathcal{F}(I_D^*[n]w[n]) \langle H[k] \rangle \right), \quad (\text{A.2})$$

so that the result of the calculation is a **windowed** characteristic LIF signal output. It is possible to divide the window out of the characteristic signal to recover the un-windowed TFLIF output signal, but the division also scales noise. This leads to an undesirable situation where SNR depends strongly on time and becomes unacceptable near the window edges. The Tukey Window is chosen for the purposes of TFLIF because it leaves a large region in the center of the window that corresponds to the characteristic output signal alone without weighting by the window function. When the Tukey window is used in this dissertation, only data within the $w(x) = 1$ region is plotted or used to calculate statistical properties. Notice that the length of time plotted in Figure A.2 decreases as r increases and the triggered average decay becomes less pronounced. The reason is that each successive plot includes a smaller central region with a later initial time where the triggered average decay has further reduced the oscillation amplitude. Also note that the plotted region begins not at the first point where $w(x) = 1$ but at the first trigger within the $w(x) = 1$ region; hence each plot starts at a maximum in the oscillation.

The effect of window settings was tested for four different Tukey windows: $r = 0$ (rectangular), 0.1, 0.25, 0.5. The example data presented here come from the cathode experiment with random oscillations of Section 6.4. Table A.1 shows that larger r values result in a smaller RMS residual between TFLIF and TALIF signals, or equivalently a higher SNR, possibly as a result of the lower side lobe levels.

Figure A.2 shows heat maps of the TFLIF signal together with the TALIF signal and the residual between them for all four example r values. In addition to the increase in random noise

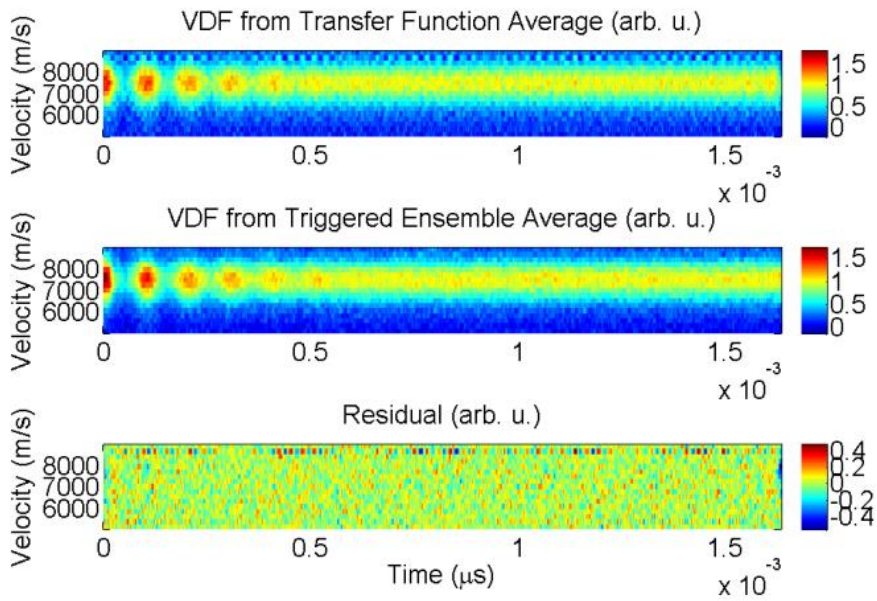
for the wider windows, the heat maps show that artifacts can also appear with the wider windows. High-frequency ringing is clearly visible in a single velocity group for the rectangular window at a higher speed than most of the distribution. The ringing is completely non-physical and can appear apparently randomly in any velocity group depending on settings such as chunk length. A different artifact is visible in the $r = 0.1$ heat maps, with a systematic discrepancy in the oscillation amplitude near the edges of the central window region. Both types of systematic error are apparently eliminated or at least reduced below the random noise level for the narrower windows.

Table A.1. The RMS residual between TFLIF and TALIF signals decreases as the r parameter of the Tukey window increases.

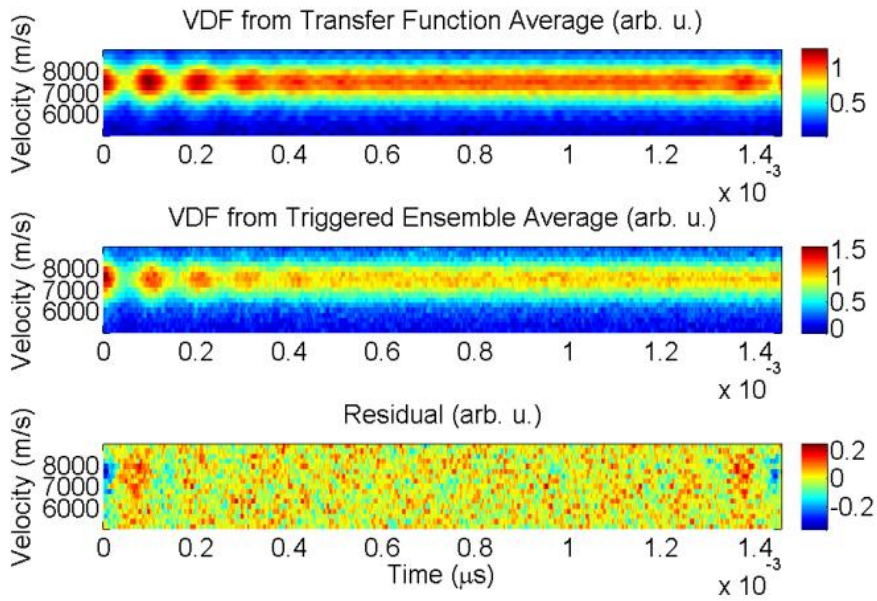
r	RMS residual
0.5	0.0568
0.25	0.0575
0.1	0.0653
0 (rectangular)	0.0978

The Tukey window with $r = 0.5$ was chosen as the standard window used for most of the analysis presented in this dissertation due to the superior characteristics for both random and systematic error. A rectangular window is sometimes used when viewing the whole chunk or time series is important, such as to detect the decay of a triggered average, but doing so increases random and systematic error. Using $r = 0.25$ could also be a reasonable choice when viewing a long time series is a priority. It would increase the presentable time series length (the central region where $w(x) = 1$) by 50% while increasing noise only slightly.

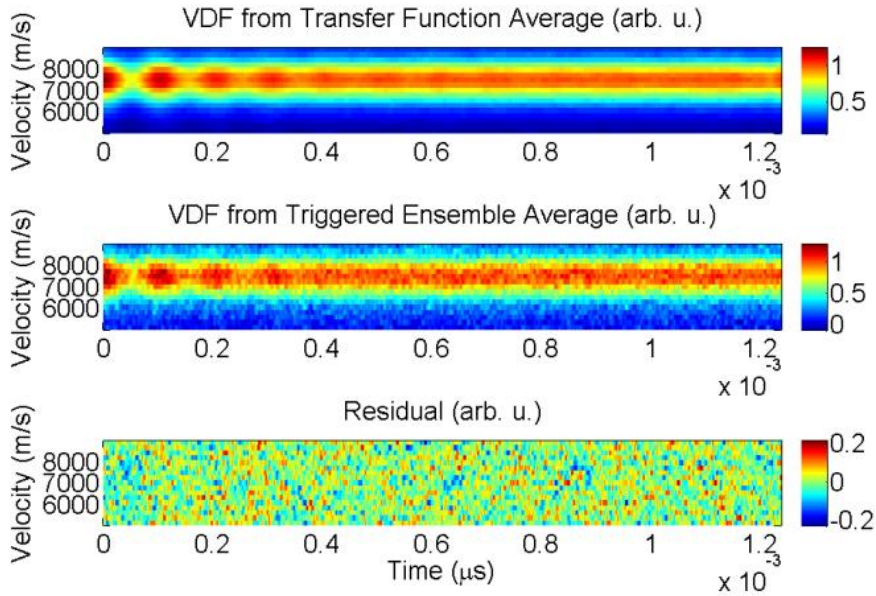
A. Rectangular window (Tukey window with $r = 0.0$):



B. Tukey window with $r = 0.1$:



C. Tukey window with $r = 0.25$:



D. Tukey window with $r = 0.5$:

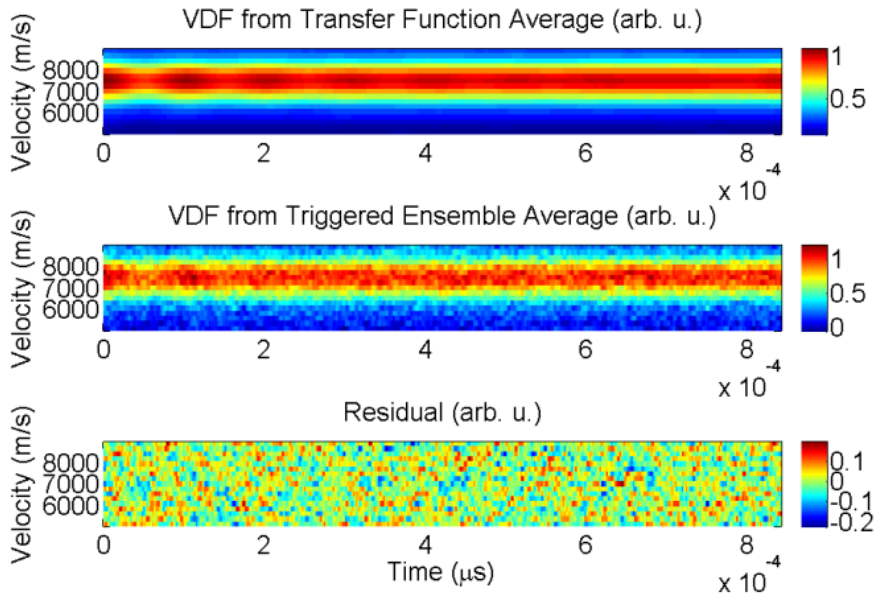


Figure A.2. Heat maps of the TFLIF signal (top), TALIF signal (middle), and their residual (bottom) for four parameter values of (A) $r = 0$, (B) $r = 0.1$, (C) $r = 0.25$, and (D) $r = 0.5$. Both random and systematic error are smaller for large r values.

Appendix B

An Analytical Model of Synchronized Photon Counting

As the most mature and widely used technique to measure TRLIF signal in Hall thrusters, we will compare the signal recovery performance of the TFLIF technique to the synchronized photon counting technique. It was developed by Mazouffre et al. and used with a number of thrusters such as the PPSX000 [79], PPS100 [80], and PPS100-ML [81], and a 200-W Hall thruster with permanent magnets [84].

We can model photon counts in time bins as Poisson random variables. A Poisson random variable is a very close approximation because the main requirement for the Poisson approximation is that there are a large number of trials that each have a small probability of success (cf. any text on probability, e.g. [106]). In this case, for either signal or noise we have a large number of ions in the interrogation zone that each have a small chance of emitting a photon that is later collected in the given time bin. Using Poisson random variables to model the number of photons collected within a time bin is a classic method to derive shot noise [107], and detecting particles/photons at low count rate can demonstrate a Poisson distribution experimentally [108]. A model based on Poisson random variables will not capture any subtleties in the noise caused by variations in the photoelectron amplitude, the action of the discriminator, Johnson noise, etc., but it will capture the effect of shot noise and background light, the dominant source of noise for LIF experiments in plasmas.

We consider both the LIF signal photons and background noise photons counted in each bin to be described by a series of Poisson random variables with each element n corresponding to the photon count of the n^{th} time bin. The number of counts collected in the n^{th} time bin after the q^{th} trigger is a function of Poisson random variables:

$$C_q[n] = \begin{cases} F_q[n] + N_q[n] & \text{if the laser is on} \\ N_q[n] & \text{if the laser is off,} \end{cases} \quad (\text{B.1})$$

where $F_q[n] \sim \text{Pois}(\lambda_f[n])$ and $N_q[n] \sim \text{Pois}(\lambda_N[n])$ are Poisson random variables describing the number of LIF and noise photons collected. $\lambda_f[n]$ is the average number of LIF photoelectron

pulses and $\lambda_N[n]$ is the average number of background noise pulses counted in the n^{th} time bin. The only parameter needed to fully specify a Poisson variable distribution is the average.

Similar to the triggered averaging technique in Subsection 4.3.4, the main implicit assumption is that essentially the same repeatable process occurs after each trigger and that fact is used to recover the average TRLIF signal waveform. A simple averaging process would not remove the background signal since the noise from background signal does not have zero mean (because photon counts are never negative). The strategy to reject noise and recover the time-correlated waveform is to add (subtract) counts taken while the laser is on (off) in sequential time bins after each trigger. The actual experiment used laser modulation at a low frequency (20 Hz) so that there are many oscillation cycles when the laser is on and then many other complete oscillation cycles when the laser is off. Since Hall thruster oscillations tend to be non-periodic, the oscillations were forced to be quasi-periodic by periodically cutting the discharge current at 2500 Hz for a short time. The time of current cutoff is used as the trigger to synchronize the addition/subtraction of photon counts with a repeatable physical process.

We model this process with a random variable $F_Q[n]$, which is the total number of counts in the n^{th} time bin after a total of Q triggers:

$$F_Q[n] = \sum_{q=1}^{Q/2} C_q[n] - \sum_{q=Q/2+1}^Q C_q[n] = \sum_{q=1}^{Q/2} F_q[n] + N_q[n] - \sum_{q=Q/2+1}^Q N_q[n]. \quad (\text{B.2})$$

For convenience, we have numbered all of the triggers taken with the laser on as 1 through $Q/2$ and those taken with the laser off as $Q/2+1$ through Q . In terms of the actual experiment, this is not chronological ordering since in general the acquisition will be over many modulation cycles. There must be at least an approximately equal number of triggers for the laser on and off cases; otherwise the background photons would not cancel appropriately.

Using the properties of sums of random variables, we see that this scheme results in a signal in the n^{th} time bin that is another random variable (whose distribution we need not calculate) with mean μ and variance σ^2 given below:

$$F_Q[n] = R(\mu = \frac{Q}{2} \lambda_f[n], \sigma^2 = \frac{Q}{2} \lambda_f[n] + Q \lambda_N[n]). \quad (\text{B.3})$$

This leads directly to an RMS SNR estimate of:

$$\text{SNR}[n] = \frac{\lambda_f[n]Q/2}{\sqrt{\lambda_f[n]Q/2 + Q\lambda_N[n]}} = \sqrt{\frac{Q}{2}} \frac{\lambda_f[n]}{\sqrt{\lambda_f[n] + 2\lambda_N[n]}}. \quad (\text{B.4})$$

It is interesting to note that this model includes contributions to noise not only from background signal but also from shot noise, so the two dominant contributions to noise in the photon counting scheme are taken into account. Consider the limiting case of no background signal with $\lambda_N[n] = 0$. Then Q is no longer meaningful, but $Q/2$ is the number of measurements capturing signal photons. The signal-to-noise ratio rises as the square root of the number of summed measurements, and the SNR for one measurement is the square root of the average photon count rate in the bin $\lambda_f[n]$, the classic prediction for shot noise.

Now, the lambdas represent the average number of photons in the n^{th} time bin, but we wish to examine the effect of varying time bin size. Assuming there is little change in the average photon collection rate as a function of time on the time scale of a bin, each lambda will simply be some average rate per second multiplied by the width of the time bin in seconds:

$$\begin{aligned}\lambda_f[n] &= R_f[n]T_B \\ \lambda_N[n] &= R_N[n]T_B,\end{aligned}\tag{B.5}$$

where R_f and R_N are the signal and noise photon rates (s^{-1}) and T_B is the width of the time bin (s). A changing collection rate inside the time bin would imply that some time-resolved behavior is being averaged out, which is undesirable, and hence the assumption of constant rates is reasonable as long as bins have been set small enough to not average out time dependent behavior. Substituting this form into the SNR equation gives:

$$SNR[n] = \sqrt{\frac{Q}{2}} \sqrt{T_B} \frac{R_f[n]}{\sqrt{R_f[n] + 2R_N[n]}}.\tag{B.6}$$

As a simple consistency check, we compare the prediction of this equation with results reported from a test using this technique [81]. Our convention is to define SNR by the time-averaged signal at the peak of the velocity distribution divided by RMS noise. Now, the average photon rates were reported to be $R_f = 10^4 \text{ s}^{-1}$ and $R_N = 10^6 \text{ s}^{-1}$, but the signal photon rate at the peak (3490 m/s) is significantly higher than the average, which leads to a poor estimate if we use the reported value of R_f . A better estimate for the average photon rate at the peak of the distribution is $R_f = 2.8 \times 10^4 \text{ s}^{-1}$, considering that the average number of signal photons collected is about 1400 over about 5×10^5 bins of size 100 ns. We then estimate that for the particular experiment reported the SNR was about:

$$SNR[n] = 20 \sqrt{\frac{Q}{2}} \sqrt{T_B}.\tag{B.7}$$

SNR was not reported numerically, but Figure 6 in that paper shows several traces of photon counts as a function of time for many velocity groups. We estimate the average signal and peak-to-peak noise by eye (ignoring the reignition transient). The RMS noise is estimated as a factor of 1/6 lower than peak-to-peak since, for a short length of data, we can expect the peak-to-peak to be about ± 3 standard deviations, which yields an estimate of $\text{SNR} = 7$ for the velocity group with 3490 m/s. Using the above equation with the estimated SNR and the time bin setting of $T_B = 100$ ns from that particular experiment gives $Q = 2.5 \times 10^6$, which is apparently an overestimate for the number of triggers used in the average, reported to be about one million, but is a reasonable estimate given the simple model and the crude estimates of the final SNR and average photon rates. The prediction of the model agrees within an order of magnitude of the experiment, providing some support that it captures the most salient features.

References

- [1] E. Choueiri, "Fundamental difference between the two variants of Hall thrusters - SPT and TAL," presented at the 37th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Salt Lake City, UT, 2001.
- [2] V. Kim, G. Popov, B. Arkhipov, V. Murashko, O. Gorshkov, A. Koroteyev, V. Garkusha, A. Semenkin, and S. Tverdokhlebov, "Electric propulsion activity in Russia," presented at the 27th International Electric Propulsion Conference, Pasadena, CA, 2001.
- [3] V. V. Zhurin, H. R. Kaufman, and R. S. Robinson, "Physics of closed drift thrusters," *Plasma Sources Sci. Technol.*, vol. 8, no. 1, p. R1, 1999.
- [4] E. Y. Choueiri, "Plasma oscillations in Hall thrusters," *Phys. Plasmas*, vol. 8, no. 4, p. 1411, 2001.
- [5] G. S. Janes and R. S. Lowder, "Anomalous Electron Diffusion and Ion Acceleration in a Low-Density Plasma," *Phys. Fluids*, vol. 9, no. 6, p. 1115, 1966.
- [6] M. McDonald, C. K. Bellant, B. A. S. Pierre, and A. D. Gallimore, "Measurement of cross-field electron current in a Hall thruster due to rotating spoke instabilities," in *Proceedings of the 47th Joint Propulsion Conference*, San Diego, California, 2011.
- [7] R. B. Lobbia and A. D. Gallimore, "High-speed dual Langmuir probe," *Rev. Sci. Instrum.*, vol. 81, no. 7, p. 73503, 2010.
- [8] M. S. McDonald and A. D. Gallimore, "Parametric Investigation of the Rotating Spoke Instability in Hall Thrusters," presented at the 32nd International Electric Propulsion Conference, Wiesbaden, Germany, 2011.
- [9] M. J. Sekerak, B. W. Longmier, A. D. Gallimore, D. L. Brown, R. R. Hofer, and J. E. Polk, "Azimuthal Spoke Propagation in Hall Effect Thrusters," presented at the 33rd International Electric Propulsion Conference, Washington, D.C., 2013.
- [10] M. J. Sekerak, R. R. Hofer, J. E. Polk, B. W. Longmier, A. Gallimore, and D. L. Brown, "Mode Transitions in Hall Effect Thrusters," in *49th AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, San Jose, CA, 2013.
- [11] A. Diallo, S. Keller, Y. Shi, Y. Raitses, and S. Mazouffre, "Time-resolved ion velocity distribution in a cylindrical Hall thruster: Heterodyne-based experiment and modeling," *Rev. Sci. Instrum.*, vol. 86, no. 3, p. 33506, Mar. 2015.
- [12] D. Goebel and I. Katz, *Fundamentals of Electric Propulsion: Ion and Hall Thrusters*. Hoboken, N.J: Wiley, 2008.
- [13] R. G. Jahn, *Physics of electric propulsion*. Mineola, N.Y: Dover Publications, Inc, 2006.
- [14] R. G. Jahn and E. Y. Choueiri, "Electric Propulsion," in *Encyclopedia of Physical Science and Technology*, Elsevier, 2003, pp. 125–141.
- [15] M. Martinez-Sanchez and J. E. Pollard, "Spacecraft electric propulsion-an overview," *J. Propuls. Power*, vol. 14, no. 5, pp. 688–699, 1998.
- [16] J. Dankanich, "Low-Thrust Mission Design and Application," presented at the 46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Nashville, TN, 2010.
- [17] A. Ketsdever, M. Young, A. Pancotti, and J. Mossman, "An Overview of Advanced Concepts for Space Access," presented at the 44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Hartford, CT, 2008.
- [18] R. Florenz, A. D. Gallimore, and P. Y. Peterson, "Developmental status of a 100-kW class laboratory nested channel Hall thruster," in *32nd International Electric Propulsion Conference, Wiesbaden, Germany*, Wiesbaden, Germany, 2011, pp. 1–9.
- [19] R. R. Hofer and A. D. Gallimore, "Efficiency analysis of a high-specific impulse Hall thruster," presented at the 40th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Ft. Lauderdale, Florida, 2004.
- [20] L. Garrigues, S. Mazouffre, and G. Bourgeois, "Computed versus measured ion velocity distribution functions in a Hall effect thruster," *J. Appl. Phys.*, vol. 111, no. 11, p. 113301, 2012.

- [21] D. M. Goebel, K. K. Jameson, R. M. Watkins, I. Katz, and I. G. Mikellides, "Hollow cathode theory and experiment. I. Plasma characterization using fast miniature scanning probes," *J. Appl. Phys.*, vol. 98, no. 11, p. 113302, 2005.
- [22] I. G. Mikellides, I. Katz, D. M. Goebel, and J. E. Polk, "Hollow cathode theory and experiment. II. A two-dimensional theoretical model of the emitter region," *J. Appl. Phys.*, vol. 98, no. 11, p. 113303, 2005.
- [23] I. G. Mikellides, I. Katz, D. M. Goebel, J. E. Polk, and K. K. Jameson, "Plasma processes inside dispenser hollow cathodes," *Phys. Plasmas*, vol. 13, no. 6, p. 63504, 2006.
- [24] D. Goebel, K. K. Jameson, and R. R. Hofer, "Hall Thruster Cathode Flow Impact on Coupling Voltage and Cathode Life," *J. Propuls. Power*, vol. 28, no. 2, pp. 355–363, Mar. 2012.
- [25] A. Hillier, R. Branam, R. Huffman, J. Szabo, and S. Paintal, "High Thrust Density Propellants in Hall Thrusters," presented at the 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, Orlando, Florida, 2011.
- [26] D. Massey, L. King, and J. Makela, "Development of a Direct Evaporation Bismuth Hall Thruster," presented at the 44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Hartford, CT, 2008.
- [27] J. Szabo, B. Pote, S. Paintal, M. Robin, G. Kolencik, A. Hillier, R. Branam, and R. Huffman, "Performance Evaluation of an Iodine Vapor Hall Thruster," presented at the 47th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, San Diego, California, 2011.
- [28] S.-C. Hwang, R. D. Lein, and D. A. Morgan, "Noble Gases," in *Kirk-Othmer Encyclopedia of Chemical Technology*, John Wiley & Sons, Inc., Ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2005.
- [29] A. Smirnov, "Plasma measurements in a 100 W cylindrical Hall thruster," *J. Appl. Phys.*, vol. 95, no. 5, p. 2283, 2004.
- [30] K. A. Polzin, T. E. Markusic, B. J. Stanojev, A. DeHoyos, Y. Raitses, A. Smirnov, and N. J. Fisch, "Performance of a Low-Power Cylindrical Hall Thruster," *J. Propuls. Power*, vol. 23, no. 4, pp. 886–888, Jul. 2007.
- [31] R. B. Lobbia and A. D. Gallimore, "Performance measurements from a cluster of four Hall thrusters," presented at the 30th International Electric Propulsion Conference, Florence, Italy, 2007.
- [32] I. Mikellides, I. Katz, and R. Hofer, "Design of a Laboratory Hall Thruster with Magnetically Shielded Channel Walls, Phase I: Numerical Simulations," presented at the 47th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, San Diego, California, 2011.
- [33] R. Hofer, D. Goebel, I. Mikellides, and I. Katz, "Design of a Laboratory Hall Thruster with Magnetically Shielded Channel Walls, Phase II: Experiments," presented at the 48th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Atlanta, GA, 2012.
- [34] D. M. Goebel, R. R. Hofer, I. G. Mikellides, I. Katz, J. E. Polk, and B. N. Dotson, "Conducting Wall Hall Thrusters," *IEEE Trans. Plasma Sci.*, pp. 1–1, 2014.
- [35] R. R. Hofer, R. S. Jankovsky, and A. D. Gallimore, "High-Specific Impulse Hall Thrusters, Part 1: Influence of Current Density and Magnetic Field," *J. Propuls. Power*, vol. 22, no. 4, pp. 721–731, Jul. 2006.
- [36] R. B. Lobbia and A. D. Gallimore, "A method of measuring transient plume properties," in *AIAA-2008-4650*.
- [37] R. B. Lobbia and A. D. Gallimore, "Fusing spatially and temporally separated single-point turbulent plasma flow measurements into two-dimensional time-resolved visualizations," in *12th International Conference on Information Fusion*, Seattle, WA, 2009, pp. 678–685.
- [38] J. B. Parker, Y. Raitses, and N. J. Fisch, "Transition in electron transport in a cylindrical Hall thruster," *Appl. Phys. Lett.*, vol. 97, no. 9, p. 91501, 2010.
- [39] C. L. Ellison, Y. Raitses, and N. J. Fisch, "Fast Camera Imaging of Hall Thruster Ignition," *IEEE Trans. Plasma Sci.*, vol. 39, no. 11, pp. 2950–2951, Nov. 2011.

- [40] M. S. McDonald and A. D. Gallimore, "Comparison of breathing and spoke mode strength in the H6 Hall thruster using high speed imaging," in *Proceedings of the 33rd International Electric Propulsion Conference*, 2013.
- [41] E. T. Dale and A. D. Gallimore, "High-Speed Image Analysis and Filtered Imaging of Nested Hall Thruster Oscillations," presented at the 34th International Electric Propulsion Conference, Hyogo-Kobe, Japan, 2015.
- [42] G. N. Tilinin, "High-frequency plasma waves in a Hall accelerator with an extended acceleration zone," *Sov Phys Tech Phys*, vol. 22, no. 8, pp. 974–978, Aug. 1977.
- [43] D. L. Brown and A. D. Gallimore, "Investigation of low discharge voltage hall thruster operating modes and ionization processes," in *31st International Electric Propulsion Conference*, Ann Arbor, Michigan, 2009.
- [44] M. J. Sekerak, "Plasma oscillations and operational modes in hall Effect Thrusters," Ph.D. dissertation, University Of Michigan, Ann Arbor, Michigan, 2014.
- [45] W. Huang, B. Drenkow, and A. D. Gallimore, "Laser-Induced Fluorescence of Singly-Charged Xenon inside a 6-kW Hall Thruster," in *45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, Denver, Colorado, 2009.
- [46] W. Hargus and M. Nakles, "Evolution of the Ion Velocity Distribution in the Near Field of the BHT-200-X3 Hall Thruster," presented at the 42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Sacramento, California, 2006.
- [47] S. Bechu, N. Gascon, S. Roche, M. Prioul, L. Albarede, P. Lasgorceix, and M. Dudeck, "Comparison between two kinds of Hall thrusters - SPT100 and ATON," presented at the 36th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Huntsville, AL, 2000.
- [48] S. Roche, S. Bechu, M. Prioul, D. Pagnon, N. Gascon, A. Bouchoule, L. Magne, P. Lasgorceix, and M. Touzeau, "Plasma characterisation of an ATON-Hall thruster - Channel and plume investigation," presented at the 36th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Huntsville, AL, 2000.
- [49] N. Yamamoto, T. Nakagawa, K. Komurasaki, and Y. Arakawa, "Influence of Discharge Oscillation on Hall Thruster Performance," presented at the 27th International Electric Propulsion Conference, Pasadena, CA, 2001.
- [50] S. Barral and E. Ahedo, "Theoretical Study of the Breathing Mode in Hall Thrusters," presented at the 42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Sacramento, California, 2006.
- [51] J. Fife, M. Martinez-Sanchez, and J. Szabo, "A numerical study of low-frequency discharge oscillations in Hall thrusters," presented at the 33rd Joint Propulsion Conference and Exhibit, Seattle, WA, 1997.
- [52] J. P. Boeuf and L. Garrigues, "Low frequency oscillations in a stationary plasma thruster," *J. Appl. Phys.*, vol. 84, no. 7, p. 3541, 1998.
- [53] N. Gascon, M. Dudeck, and S. Barral, "Wall material effects in stationary plasma thrusters. I. Parametric studies of an SPT-100," *Phys. Plasmas*, vol. 10, no. 10, p. 4123, 2003.
- [54] S. Barral, E. Ahedo, H.-J. Hartfuss, M. Dudeck, J. Musielok, and M. J. Sadowski, "On the Origin of Low Frequency Oscillations in Hall Thrusters," in *AIP Conf. Proc.*, 2008, vol. 993, pp. 439–442.
- [55] K. Hara, M. J. Sekerak, I. D. Boyd, and A. D. Gallimore, "Mode transition of a Hall thruster discharge plasma," *J. Appl. Phys.*, vol. 115, no. 20, p. 203304, May 2014.
- [56] K. Hara, M. J. Sekerak, I. D. Boyd, and A. D. Gallimore, "Perturbation analysis of ionization oscillations in Hall effect thrusters," *Phys. Plasmas*, vol. 21, no. 12, p. 122103, Dec. 2014.
- [57] K. Hara, I. D. Boyd, M. J. Sekerak, and A. D. Gallimore, "Breathing mode in Hall effect thrusters," presented at the 34th International Electric Propulsion Conference, Hyogo-Kobe, Japan, 2015.
- [58] R. B. Lobbia, M. J. Sekerak, R. Liang, and A. D. Gallimore, "High-speed Dual Langmuir Probe Measurements of the Plasma Properties and EEDFs in a HET plume," presented at the 32nd International Electric Propulsion Conference, Wiesbaden, Germany, 2011.

- [59] J. E. Hansen and W. Persson, "Revised analysis of singly ionized xenon, Xe II," *Phys. Scr.*, vol. 36, no. 4, pp. 602–643, Oct. 1987.
- [60] L. Broström, S. Mannervik, A. Passian, and G. Sundström, "Investigation of some transitions and lifetimes in Xe i i ," *Phys. Rev. A*, vol. 49, no. 5, p. 3333, 1994.
- [61] V. N. Očkin and S. Kittell, *Spectroscopy of low temperature plasma*, 1. Aufl. Weinheim: WILEY-VCH, 2009.
- [62] T. B. Smith, "Deconvolution of ion velocity distributions from laser-induced fluorescence spectra of xenon electrostatic thruster plumes," Ph.D. dissertation, University Of Michigan, Ann Arbor, Michigan, 2003.
- [63] N. Sadeghi, "Molecular Spectroscopy Techniques Applied for Processing Plasma Diagnostics," *J Plasma Fusion Res*, vol. 80, no. 9, pp. 767–776, 2004.
- [64] S. Mazouffre, "Laser-induced fluorescence diagnostics of the cross-field discharge of Hall thrusters," *Plasma Sources Sci. Technol.*, vol. 22, no. 1, p. 13001, Feb. 2013.
- [65] W. Huang, T. Smith, and A. Gallimore, "Obtaining Velocity Distribution Using a Xenon Ion Line with Unknown Hyperfine Constants," presented at the 40th AIAA Plasmadynamics and Laser Conference, San Antonio, 2009.
- [66] U. Fantz, "Basics of plasma spectroscopy," *Plasma Sources Sci. Technol.*, vol. 15, no. 4, pp. S137–S147, Nov. 2006.
- [67] E. Pawelec, S. Mazouffre, and N. Sadeghi, "Hyperfine structure of some near-infrared Xe I and Xe II lines," *Spectrochim. Acta Part B At. Spectrosc.*, vol. 66, no. 6, pp. 470–475, Jun. 2011.
- [68] W. Huang, "Study of Hall thruster discharge channel wall erosion via optical diagnostics," Ph.D. dissertation, University Of Michigan, Ann Arbor, Michigan, 2011.
- [69] B. M. Reid, "The influence of neutral flow rate in the operation of Hall thrusters," Ph.D. dissertation, University Of Michigan, Ann Arbor, Michigan, 2009.
- [70] H.-J. Kunze, *Introduction to plasma spectroscopy*. Berlin ; New York: Springer, 2009.
- [71] W. A. Hargus, "Laser-Induced-Fluorescence-Derived Hall Effect Thruster Ion Velocity Distribution Visualization," *IEEE Trans. Plasma Sci.*, vol. 39, no. 11, pp. 2918–2919, Nov. 2011.
- [72] P.-Q. Elias, J. Jarrige, E. Cucchetti, D. Packan, and A. Bulit, "Full Ion Velocity Distribution Function measurement in an Electric Thruster using LIF-based tomographic reconstruction.pdf," presented at the 34th International Electric Propulsion Conference, Hyogo-Kobe, Japan, 2015.
- [73] S. Svanberg, *Atomic and molecular spectroscopy: basic aspects and practical applications*, 4th, ed ed. Berlin ; New York: Springer, 2004.
- [74] E. L. Hill, "Relative Intensities in Nuclear Spin Multiplets," *Proc. Natl. Acad. Sci.*, vol. 15, no. 10, pp. 779–784, 1929.
- [75] A. E. Siegman, *Lasers*. Mill Valley, Calif: Univ. Science Books, 1986.
- [76] W. A. Hargus and M. R. Nakles, "Ion Velocity Measurements Within the Acceleration Channel of a Low-Power Hall Thruster," *IEEE Trans. Plasma Sci.*, vol. 36, no. 5, pp. 1989–1997, Oct. 2008.
- [77] E. Scime, C. Biloiu, C. Compton, F. Doss, D. Venture, J. Heard, E. Choueiri, and R. Spektor, "Laser induced fluorescence in a pulsed argon plasma," *Rev. Sci. Instrum.*, vol. 76, no. 2, p. 26107, 2005.
- [78] I. A. Biloiu, X. Sun, and E. E. Scime, "High time resolution laser induced fluorescence in pulsed argon plasma," *Rev. Sci. Instrum.*, vol. 77, no. 10, p. 10F301, 2006.
- [79] S. Mazouffre, D. Gawron, and N. Sadeghi, "A time-resolved laser induced fluorescence study on the ion velocity distribution function in a Hall thruster after a fast current disruption," *Phys. Plasmas*, vol. 16, no. 4, p. 43504, 2009.
- [80] G. Bourgeois, S. Mazouffre, and N. Sadeghi, "Examination of the temporal characteristics of the electric field in a Hall effect thruster using a photon-counting technique," in *31st International Electric Propulsion Conference, Ann Arbor, MI*, 2009.
- [81] S. Mazouffre and G. Bourgeois, "Spatio-temporal characteristics of ion velocity in a Hall thruster discharge," *Plasma Sources Sci. Technol.*, vol. 19, no. 6, p. 65018, Dec. 2010.

- [82] J. Vaudolon, L. Balika, and S. Mazouffre, "Photon counting technique applied to time-resolved laser-induced fluorescence measurements on a stabilized discharge," *Rev. Sci. Instrum.*, vol. 84, no. 7, p. 73512, 2013.
- [83] J. Vaudolon, B. Khiar, and S. Mazouffre, "Low and high frequency oscillations of the accelerating electric field in a Hall thruster," in *Proceedings of the 33rd International Electric Propulsion Conference, Washington, DC*, 2013.
- [84] J. Vaudolon and S. Mazouffre, "Investigation of the ion transit time instability in a Hall thruster combining time-resolved LIF spectroscopy and analytical calculations," in *34th International Electric Propulsion Conference, Hyogo-Kobe, Japan*, 2015.
- [85] A. Diallo, Y. Shi, S. Keller, Y. Raitses, and S. Mazouffre, "Time-Dependent Ion Velocity Distribution: A novel Heterodyne Laser-Induced Fluorescence with Coupled Wave Excitation," presented at the 33rd International Electric Propulsion Conference, Washington, D.C., 2013.
- [86] N. A. MacDonald, M. A. Cappelli, and W. A. Hargus, "Time-synchronized continuous wave laser-induced fluorescence on an oscillatory xenon discharge," *Rev. Sci. Instrum.*, vol. 83, no. 11, p. 113506, 2012.
- [87] N. MacDonald-Tenenbaum, C. V. Young, A. L. Fabris, M. Nakles, W. Hargus Jr, and M. A. Cappelli, "Time-Synchronized Continuous Wave Laser Induced Fluorescence Velocity Measurements of a 600 Watt Hall Thruster," presented at the 34th International Electric Propulsion Conference, Hyogo-Kobe, Japan, 2015, vol. IEPC-2015-350.
- [88] C. Eichhorn, S. Lohle, S. Fasoulas, H. Leiter, and M. Auweter-Kurtz, "Two-Photon Spectroscopy on Neutral Xenon in the Plume of the Radio-Frequency Ion Thruster RIT-10," presented at the 32nd International Electric Propulsion Conference, Wiesbaden, Germany, 2011.
- [89] C. J. Durot, A. D. Gallimore, and T. B. Smith, "Validation and evaluation of a novel time-resolved laser-induced fluorescence technique," *Rev. Sci. Instrum.*, vol. 85, no. 1, p. 13508, 2014.
- [90] M. L. Meade, *Lock-in amplifiers: principles and applications*. London: P. Peregrinus on behalf of the Institution of Electrical Engineers, 1983.
- [91] P. M. T. Broersen, "Mean square error of the Empirical Transfer Function Estimator for stochastic input signals," *Automatica*, vol. 40, no. 1, pp. 95–100, Jan. 2004.
- [92] W. P. Heath, "Choice of weighting for averaged non-parametric transfer function estimates," in *43rd IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas, 2004.
- [93] P. Guillaume, I. Kollar, and R. Pintelon, "Statistical analysis of nonparametric transfer function estimates," *Instrum. Meas. IEEE Trans. On*, vol. 45, no. 2, pp. 594–600, 1996.
- [94] W. R. Bennett, "Spectra of Quantized Signals," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 446–472, Jul. 1948.
- [95] P. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms," *IEEE Trans. AUDIO ELECTROACOUSTICS*, vol. 15, no. 2.
- [96] M. S. Bartlett, "Periodogram analysis and continuous spectra," *Biometrika*, vol. 37, no. 1–2, pp. 1–16, 1950.
- [97] J. W. Cooley, P. A. W. Lewis, and P. D. Welch, "Historical notes on the fast Fourier transform," *Proc. IEEE*, vol. 55, no. 10, pp. 1675–1677, 1967.
- [98] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, 1978.
- [99] P. Y. Peterson and A. D. Gallimore, "The performance and plume characterization of a laboratory gridless ion thruster with closed electron drift acceleration," presented at the 40th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Fort Lauderdale, Florida, 2004.
- [100] W. Huang, B. M. Reid, T. B. Smith, and A. D. Gallimore, "Laser-Induced Fluorescence of Singly-Charged Xenon in a 6kW Hall Thruster," presented at the 44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Hartford, CT, 2008.

- [101] M. Georgan, C. Durot, and A. D. Gallimore, "Preliminary Measurements of Time-Resolved Ion Velocity Distributions Near a Hollow Cathode," presented at the 34th International Electric Propulsion Conference, Hyogo-Kobe, Japan, 2015.
- [102] S. . Kaufman, "HIGH-RESOLUTION LASER SPECTROSCOPY IN FAST BEAMS," *Opt. Commun.*, vol. 17, no. 3, pp. 309–312, Jun. 1976.
- [103] D. Field and M. D. Gray, "Kinematic compression and expansion of the velocity distributions of particles in gas flows," *Phys. Rev. A*, vol. 40, no. 4, p. 1976, 1989.
- [104] J. Pérez-Luna, G. J. M. Hagelaar, L. Garrigues, and J. P. Boeuf, "Method to obtain the electric field and the ionization frequency from laser induced fluorescence measurements," *Plasma Sources Sci. Technol.*, vol. 18, no. 3, p. 34008, Aug. 2009.
- [105] R. Spektor, "Computation of two-dimensional electric field from the ion laser induced fluorescence measurements," *Phys. Plasmas*, vol. 17, no. 9, p. 93503, 2010.
- [106] S. M. Ross, *A first course in probability*, 8th ed. Upper Saddle River, N.J: Pearson Prentice Hall, 2010.
- [107] M. Johnson, *Photodetection and measurement maximizing performance in optical systems*. New York: McGraw-Hill, 2003.
- [108] W. T. Vetterling, "Demonstrating Poisson statistics," *Am. J. Phys.*, vol. 48, no. 5, p. 342, 1980.