REAL-TIME STATE ESTIMATION IN PLASMA MODELING APPLICATIONS

A Dissertation

by

CHRISTINE MARIE GREVE

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Manoranjan Majji
Co-Chair of Committee,	Kentaro Hara
Committee Members,	Richard Miles
	Douglas Allaire
Head of Department,	Ivett Leyva

August 2022

Major Subject: Aerospace Engineering

ABSTRACT

Development of truly predictive models for plasma physics phenomena continues to pose a significant challenge to the research community. Recent interest in data-driven modeling and data assimilation have arisen in the plasma physics community to provide an alternative to predictive models or to directly solve for uncertain physics. The focus of this doctoral research is the development of a state estimation technique that uses experimental measurements to improve plasma physics models by either improving the solutions of lower-fidelity, faster running models or by estimating unknown or uncertain physics. This dissertation demonstrates that the simple class of Kalman filtering can provide significant insight to plasma modeling. Test cases begin with the canonical Lorenz chaotic attractor and a driven-damped harmonic oscillator to demonstrate the fidelity of the estimation technique as increasingly sparse measurement data are used. Then, the EKF is applied to global plasma models to demonstrate that physical states including the electron temperature, absorbed electron power, and reaction rate coefficients can be estimated with physical relevance. Additionally, test cases including complex models, measurement signals relating to multiple states, and multiple estimates being sought, simultaneously, are examined. Finally, this dissertation extends the EKF into a single spatial dimension. After two general test cases are used to demonstrate how the filter can be applied in one spatial dimension for representative cases of drift and diffusion processes, the conclusion of the dissertation focuses on the challenges of applying the EKF to a one-dimensional fluid model of a Hall effect thruster to study the anomalous component of electron mobility.

DEDICATION

For anyone who has ever looked up at the night sky and dreamt.

ACKNOWLEDGMENTS

I must first thank my advisor, Dr. Ken Hara, for an incredible amount of support throughout my doctoral study. We have accomplished no small feat together, from working to establish a research group and handling a long-distance mentorship to stepping outside of both our comfort zones to develop a data-driven model and finally graduating your first doctoral student. I am grateful for your seemingly unending patience as we waded through a new area of research for both of us, as I faced the stress of graduate school, and all the times I felt lost and uncertain of my goals and my accomplishments. Thank you for all of the breakfast conversations about the human side of research and our lab group. You have pulled me out of my determined independence streak more than once, and reminded me that asking for help is not admitting defeat.

I also want to thank my committee members. Dr. Manoranjan Majji was instrumental in the development of this technique, and has served as the most supportive and helpful co-advisor that a student could want. Taking on a student that is not fully yours can be a difficult balancing act, but Dr. Majji handled it with grace and an open mind. Thank you for teaching me the world of state estimation. I want to thank Dr. Richard Miles for adding a much-needed experimentalist perspective to my research, helping me to remember how to translate ideas and capabilities between the computational and experimental world. I also want to thank Dr. Douglas Allaire who assisted in expanding my knowledge of the world of state estimation and some of the pertinent research that has been performed.

I was fortunate to spend my first summer in graduate school at the Air Force Research Laboratory at Edwards Air Force Base in California. Those ten weeks were more instrumental to my career than I originally imagined. Under the advisement of Dr. Justin Koo, Dr. Dan Eckhardt, Dr. David Bilyeu, and Dr. Rob Martin, I was introduced to the idea of data-driven modeling for plasma physics. We spent the summer and subsequent year and a half developing a phase-space representation of computational and experimental data to compare the generalized regions of space where the dynamics of a system occur. Without this co-op, I would not have written this disserta-

tion, and I also would not be continuing similar work through a NRC Fellowship position with the same group this coming year. The desert may not be my favorite terrain, but the relationships and research make it worthwhile.

I must also thank the man who gave me my start in electric propulsion, Dr. Kurt Polzin. Kurt, we have known each other for some eight years now and I still feel as though I can come to you with any question or joke or problem. You were my first true mentor and I am so grateful for all of the advice, wisdom, and laughter you have shared with me. For all my love of space, that first internship with you at NASA MSFC really did introduce me to plasma propulsion and set me on the path that has led to this dissertation. If you ever make that startup on the beach, you know I will be first in line to work for you.

For my friends, I cannot thank each and every one of you enough. For the lab members who were with me at the start, Adnan and Rupali, we have come so far from where we started. You have been a source of motivation and sympathy these last five years as we undertook confusing courses and tried to make sense of our advisor. I also thank all of my lab members from Stanford: Andrew, Andy, Shige, Daniel, Derek, and Raymond. Though I was not on campus most of the time, you all made me feel like part of the group and were quick to welcome me anytime I made it to California. To the other students at TAMU that I almost entirely met before we even started graduate school, I would not have made it through quals or the years of graduate school without your support, humor, camaraderie, and all the crazy adventures. Hunter, Jepp, Carl, Ethan, Laura, Tyler, Johnnie, and Nooner, these last five years would not have been the same without the video games, movie nights, cookouts, study sessions, music, and rock climbing. To my girls, McKenzie, MinGin, Abby, and Traci, I am so grateful to have met you. Even more, I am blessed to have met you all while dancing. To be able to watch so many strong women go out and make their marks on the world in everything from STEM to the arts has been so motivational. To my D&D ladies, thank you for letting me decompress by throwing all manner of chaos at you this past year. Creating a world with you and for you has been such a blessing, and I hope you see my love in every corner of the world that you discover.

I must separately talk about my time dancing, both as a student and as a teacher. I was welcomed with open arms and warm hearts by teachers and fellow students alike, and quickly found myself teaching, something I had sworn I would never do. Jenny, who knew that four years ago as you sat with your youngest daughter in the back playroom of a dance studio that we would end up sitting and talking backstage at recital four years later while you held your newborn son. You took a chance on me and gave me the space and opportunity to grow and learn how to become the teacher that I am today. Even more, you have supported me and consoled me nearly every day of the last four years, knowing all too well the struggles I have faced as your husband went through similar challenges. I will never be able to thank you enough for helping me learn that I can make a difference in others and that I can create something beautiful. Ms. Karen, you are the world's greatest cheerleader and sweetest woman who has always looked out for me and made sure that I felt at home. Suma and Scott, these past few years of learning West Coast Swing from you have been not only a much-needed distraction from research, but a source of support and strength from a couple who know what it takes to be a woman in STEM. Your advice and encouragement have been invaluable. Thank you all for giving me places to call home while I was in Texas.

To my students: It has been an absolute gift to get to know each and every one of you beyond just the dancers I see in class. Watching you all grow and work hard kept me motivated to do the same in teaching and in research. Classes were usually far from perfect, and sometimes not even truly ballet, but you took all of it with open minds and excitement. You helped me push myself every week, and always understood when I made a mistake or simply could not say. a coherent sentence because I was so distracted by research. I must especially thank Karsen and Sofia. You two were some of the first students I truly got to know, and you have been with me every step of this journey. Thank you for your excitement, your strength, your personalities, your support, and the absolutely undeserved blessing of your admiration. You helped me believe that I could be a teacher and make a difference in someone's life.

I have intentionally kept you until last, Mom and Dad. I will never be able to thank you for everything you have done and sacrificed to help me reach this point. You had to trust me so many

times when I insisted that I could do this, even when I was stressed or felt there was no way forward. You have always believed in me and reminded me of how capable I am, especially those days it felt like few others believed the same. We have leaned on each other a lot these past five years for so many different reasons, and I am so proud to be your daughter. Thank you also to all of my family, grandparents, aunts, uncles, cousins, and family friends who have supported me, prayed for me, and wished me well all these years.

Lastly, and most importantly, thank you to God for all the blessings You have bestowed upon me and guidance You have given me as I navigated these five years.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of advisors Professor Ken Hara of the Aeronautics and Astronautics Department at Stanford University and Professor Manoranjan Majji as well as Professor Richard Miles of the Department of Aerospace Engineering and Professor Douglass Allaire of the Department of Mechanical Engineering.

All work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University and a dissertation research fellowship from the Department of Defense through the National Defense Science and Engineering Graduate Fellowship program.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION AND LITERATURE REVIEW

The beginning is perhaps more difficult than anything else, but keep heart...

- Vincent Van Gogh, *The Hague, Jan. 1837*

Plasma physics exists across natural phenomena and engineering devices, from lightning strikes to spaceflight. As the fourth state of matter, plasmas consist of superheated, ionized gases that exist in a variety of conditions, from high temperature fusion reactor plasmas (>100 eV) to low temperature material processing (<10 eV). High temperature plasmas are largely studied on the scale of astrophysics and fusion. More commonly, high temperature plasmas exist in the solar core [5] and nuclear fusion reactions [6]. On the lower temperature scales, plasma can be used for materials processing [7], low earth orbit propulsion and station-keeping [8], and long-duration satellite propulsion systems [9]. Low temperature plasma research began with the development of light bulbs, filaments, and probing equipment. This translates to the modern day usage of neon signs, and display screens such as televisions as well as applications in the semiconductor industry [10] and medical fields [11]. Regardless of begin high or low temperature, the resulting plasma is an amalgamation of phenomena occurring across a spectrum of timescales and length scales. Furthermore, the plasma contains both electromagnetic and gas characteristic properties that provide crucial insight to the behavior of the plasma. Such complex physics play into a variety of undesired and poorly understood phenomena across plasma applications that hinder the ability to develop predictive models of such behavior.

1.1 Plasma Processing

Plasmas are used across the materials and processing industry for a multitude of applications including microelectronics [12], plasma surface treatments [13], and film deposition [14]. Plasmas have been used to scrub carbon fiber nanotubes for water filtration [15], they can etch sensitive

Part of this chapter is reprinted with permission from "A data-driven approach to model calibration for nonlinear dynamical systems" by C. M. Greve, K. Hara, R. S. Martin, D. Q. Eckhardt, and J. W. Koo, 2019, *Journal of Applied Physics*, 125, 244901, COPYRIGHT 2019 by AIP Publishing.

materials for the semiconductor industry [10], and they can be used for processing and bathing of materials [16]. These plasmas are commonly generated as inductively coupled plasmas (ICPs) or capacitively coupled plasmas (CCPs). Regardless of the method of plasma generation, plasma processing is the result of a careful balance between plasma chemistry, plasma physics, and surface chemistry. Some experimental results can be obtained using Langmuir probes and more intricate laser-induced fluorescence, but these diagnostics equipment are limited in scope and resolution [17].

Inductively coupled plasmas use electromagnetic induction to generate oscillating magnetic fields that interact with initiated ions and electrons to produce further ionization in suitable gases. They are considered a type of radio frequency discharge, operating on scales of 0.1 to 100 MHz and exist in three configurations, planar, cylindrical, and half-toroidal. The electrode exists as a length of flat metal wound in a coil, a helical spring, or a toroidal solenoid cut along the major diameter into two equal portions, depending on the configuration. A time-varying electric field is supplied to the coil to induce a time-varying magnetic field [18]. ICPs are often found in detection systems such as mass spectrometers and are low pressure plasma sources that operate continuously or in modulation [7]. The scalability of such systems allows for wide experimental testing, but suffers from high density, stray capacitive coupling, and non-uniformity [19]. Conversely, the ability to place the electrodes outside of the reaction chamber creates discharges that are relatively free of contamination.

Capacitively coupled plasmas operate when two electrodes are placed parallel within a chamber, with one electrode grounded and the other attached to a radio-frequency (RF) power supply. The power supply creates an electric field between the electrodes, ionizing atoms and releasing electrons that are accelerated by the RF field and help to further ionize the gas. Substrates can be placed on either plate for etching or the plasma can be used as a precursor for the deposition of a thin film. The ions accelerated across the plasma sheath help to drive chemical reactions between radical particles and surface materials which form stable chemicals [20]. Due to the nonlinearity in CCPs, the sheath structures are more complex than those of traditional DC sheaths and drive the

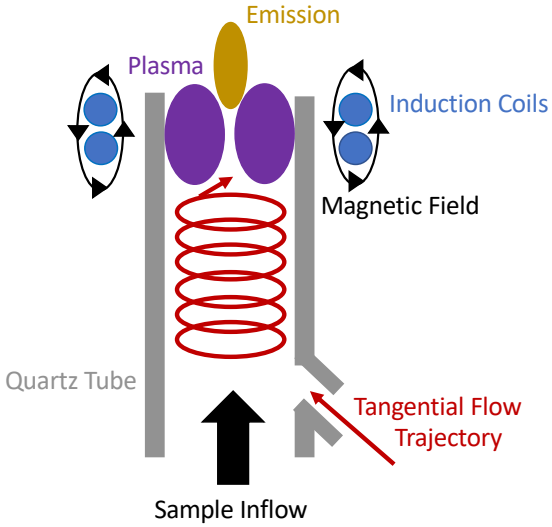

Figure 1.1: An inductively coupled plasma schematic.

development of analytical models to understand the dynamics of these sheaths [10]. These plasmas suffer from a lack of independent control over ion flux and energy, uniformity problems at high density, edge effects, and local plasma resonances.

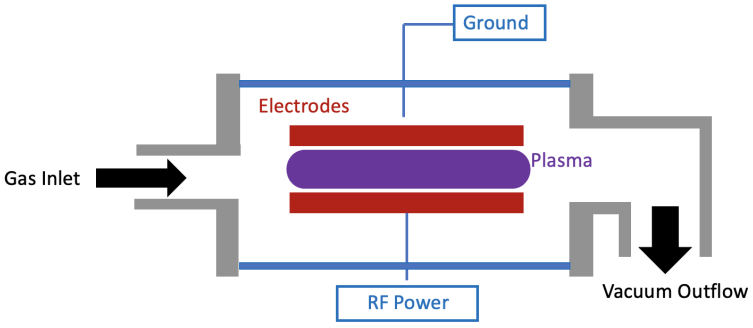

Figure 1.2: A capacitively coupled plasma schematic.

While both ICP and CCP processes have been used extensively for the past few decades, a variety of challenges still face both technologies. In many instances, using either technique for ion etching is still not refined for use with various bulk materials or complex multilayer structures. The

complex gas mixtures used in a variety of etching processes can lead to unevenness in the formation of trenches due to secondary etching [21]. Additionally, similar processes can result in uneven wall structures of etched profiles or even tips of microtrenches where locally high electric fields can accumulate. Other sources of concern with such types of industrial plasmas lie in plasma hysteresis during pulsed operation and mode transitions during plasma operation as well as the effects of electron avalanche which occur when a strong enough electric field causes indirect ionization by collisions, producing an increasing number of secondary electrons [22]. Typical operating pressures range from 100 mTorr to 1 Torr while frequencies operate on the scale of 1MHz supplied by voltages of 10 V to 1 kV. These plasmas can transition to a collisionless mode, where the exact mechanisms involved in such electron heating remain unknown [23]. Differential charging of etched materials is also a major concern as such effects can deviate ion trajectories and distort the desired etched profile [24]. Coupling between sheath thickness and applied power increases the complexity of changing incident ion energy.

1.2 Plasma Propulsion

Plasma is also at the forefront of spacecraft propulsion technologies. These plasmas span the entire range of temperatures, from the low temperature pulsed plasma thrusters and micropropulsion units such as electrosprays to the high temperature fission/fusion devices including magnetoplasmadynamics and nuclear propulsion [8, 25]. Plasma-based propulsion can fit into electrothermal systems that employ electrical currents or radiation to directly heat propellant, electrostatic systems that accelerate ionized propellant through an electric field without magnetization of ions, and electromagnetic systems that use electromagnetic forces to accelerate plasmas. So-called electric propulsion is any device that accelerates a gas by electronically heating it or through the application of electromagnetic forces. Unlike chemical propulsion, whose energy is created from the chemical reaction occurring in the combustion chamber, electric propulsion systems rely on energy from external power sources including solar panels, nuclear reactors, and radioisotope thermoelectric generators [26]. These thrusters have created great interest in the propulsion community due to their lesser energy limitation resulting from the use of such external power sources

as well as their improved propulsion performance such as specific impulse. Thus, these thrusters are ideal for missions requiring strict control over thrust profiles or those that travel distances that allow for continuous operation of the thruster to reach high velocities.

Some of the most significant parameters of space propulsion include the thrust output and specific impulse. Specific impulse is characterized by the exhaust speed of the propellant divided by the gravitational acceleration constant, g. This can be alternatively written as engine thrust divided by propellant mass flow rate. Chemical propulsion devices commonly used to escape a planet's gravitational field can produce high thrust levels, but typically have specific impulses limited to $<$ 500 s due to the high propellant flow rates and the maximum temperature of the combustion reaction. In comparison, electric propulsion can operate with specific impulse values ranging 1000 - 10,000 s, which reduces fuel consumption and consequently increases the payload fraction, e.g. the ratio of the weight of payload to the launch weight of the spacecraft. The high specific impulse enables these thrusters to be used for events such as interplanetary missions, station-keeping of satellites in orbit, and orbit raising maneuvers but the low thrust requires the thrusters to operate in excess of 10,000 hours.

The Hall effect thruster (HET) is one of the most well-known plasma propulsion devices used in modern in-space flight operations. These thrusters operate with high specific impulses ($\approx$1600 s) and better efficiencies than many other electric propulsion devices using pure noble gas propellants. Hall effect thrusters are of great advantage to space technologies due to their variable thrust capabilities, long thruster operations, low propellant mass, and assumed scalability for different mission objectives. These devices have seen thousands of flights in low earth orbit, most recently with the fleet of SpaceX Starlink satellites and the successful ThrustMe flight of an iodine-based system. Beyond the confines of low earth orbit, a gridded ion thruster, a neighbor to the HET, was used on the NASA DAWN mission while a true HET was used on the MARCO satellites that flew past Mars with the Perseverance rover. A Hall effect thruster is planned to be used on the upcoming Psyche mission by NASA. Additionally, the upcoming Lunar Gateway mission will use a set of HETs on the Propulsion and Power Unit.

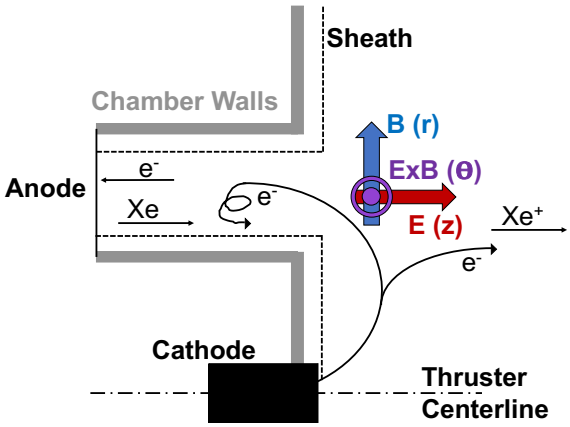

Figure 1.3: A Hall effect thruster schematic.

A simple schematic of such a thruster is provided in Fig. 1.3 that shows the direction of the electric and magnetic fields, resulting $E \times B$ azimuthal direction, and main components of the thruster. The thruster is comprised of an annular channel with either an externally or internally mounted cathode; in this case, the cathode is internally mounted. As seen in the schematic, HETs consist of an anode which supplies a positive potential and is commiserate with the location of propellant injection, a cathode that emits electrons both to neutralize the plasma as it exits the thruster channel and to supply electrons to the channel, and electrically charged coils to generate the magnetic field. A radial magnetic field is applied in the annular channel, which reduces the Larmor radius of the electrons in the channel. As the electrons have a long mean free path, the resulting azimuthal drift is generated, effectively trapping the electrons in the channel and maintaining an electric field in the axial direction. Thus, injected neutral gas collides with the trapped electrons, ionizing it in the channel and accelerating the resulting ions due to the potential gradient. Multiple variations of the HET exist based on channel material and channel length depending on the manufacturer and mission requirements.

As the future of spaceflight continues to grow, the need for improvements to thruster performance, scaling of thrusters, and increased lifespans drive the goals of the research community. As previously mentioned, the complex nature of plasma physics creates challenges to reaching these goals. The wide range of time and length scales in the particle behaviors are also present across

the variety of plasma oscillations occurring in the thruster. While some of these oscillations have been observed experimentally, such as low-frequency breathing modes and azimuthally rotating spokes, the underlying causes of these oscillations are more challenging to resolve [27, 28, 29]. This is in part due to the limitations of current experimental diagnostic tools and in part due to a lack of understanding of the underlying phenomena. Furthermore, electron dynamics operate at such small time scales that they are nearly impossible to track experimentally with the state-of-the-art diagnostic technologies, yet are crucial to the operation of these thrusters. Known phenomena includes wall collisions [3, 30, 29, 31], the resulting secondary electron emission (SEE) from such collisions [30, 31], and inexplicably high levels of diffusion in the presence of magnetic fields [32, 33, 34, 35]. These processes are notable because they lead to one of the common lifespan limiters: erosion of the thruster channel to the point where the magnetic coils and pole pieces are exposed to the plasma. Recent advances by NASA Jet Propulsion Laboratory (JPL) have created a magnetically shielded HET whose magnetic field lines are created in such a way to prevent such high levels electron interaction with the wall [36, 37]. Nonetheless, the dynamics of the electrons remain uncertain.

Because of the staunch time commitment required to physically test a thruster for its entire lifespan, computational models have been designed to study thruster performance with the hope of both extrapolating the results to determine the thruster lifespan and failure modes while also attempting to develop truly predictive models that can shape the design and scaling of future thrusters. The kinetic, small-scale nature of plasma transport and dynamics need to be better understood to correctly characterize the behavior of these plasmas. Furthermore, the behavior of these plasma devices operating in space compared to in ground test facilities must be better characterized to gain a thorough understanding of the effects of test facilities. The main focus of this dissertation will be towards improving Hall effect thruster simulations.

1.3 Plasma Modeling

Modeling enables the investigation of phenomena and locations where experimental testing is difficult or impossible, less expensive trade studies for new concepts and designs, and further

understanding of the physical phenomena that may not be able to be experimentally measured. The challenge of these models is how to represent phenomena that is poorly understood or simply unknown to the community in such a way as to improve understanding.

The complexity of plasma creates numerous opportunities to study interesting physical phenomena. Models of increasing fidelity and complexity have been created to study the different effects of plasma behavior in continued attempts to further understanding of plasma physics. As an example, the number of dimensions studied in a given computational model affects the types of studies being performed. Global models serve to track volume-averaged quantities for quick results and qualitative trends rather than quantitative solutions [38, 39, 40]. One-dimensional models have given insight into electron mobility parameters as they change across domains [31, 41]. Two-dimensional models serve as the current standard in state-of-the-art Hall effect thruster models, as they provide some of the most detailed solutions and have been used to study electron drift instabilities [42, 43, 44, 45]. Three-dimensional models have been pursued to study plume interactions and particle transport, but are so computationally expensive they have not become common to date [46, 47].

Similarly, the mathematical description of the plasma physics can affect the phenomena being studied. Different numerical approaches, such as particle-based or a fluid continuum, can be used to study different time and length scales within plasma processes. A select sample of mathematical descriptions are provided with brief descriptions below:

- Particle-in-Cell (PIC): These models trace the equations of motion for each macroparticle being followed in the simulation as

$$\frac{d\boldsymbol{x}_j}{dt} = \boldsymbol{v}_j, \tag{1.1a}$$

$$\frac{d\boldsymbol{v}_j}{dt} = \boldsymbol{a}_j, \tag{1.1b}$$

  where each particle j can use the time derivative of its position $\boldsymbol{x}$ to obtain its velocity $\boldsymbol{v}$

and similarly use the velocity derivative to obtain the acceleration $\boldsymbol{a}$. Common techniques use a leapfrog propagation method (second order) to update the locations and velocities without numerical errors [48]. Macroscopic quantities are obtained by sampling, though their use can lead to noise in the simulation due to the use of discrete numbers of particles to represent more continuous behavior [8]. Such numerical effects can be reduced by increasing the number of macroparticles, but this also increases the computational cost and does not entirely remove the effects of statistical noise. Additionally, the collisions occurring within the plasma are tracked as probabilities within each set of equations of motion. These particle-based models are used to capture nonequilibrium effects in a plasma.

- Grid-Based Direct Kinetic (DK): This method follows the ionization and charge exchange in discretized phase space where the kinetic equations are solved directly rather than by the equations of motion. The use of discretized velocity distribution functions to calculate collision integrals eliminates the statistical noise found in the PIC simulations. This method uses collisionless Boltzmann equations and common computational fluid dynamics (CFD) advection methods for propagation such as finite difference schemes and Weighted Essentially Non-Oscillatory (WENO) schemes [49]. Work by Hara uses finite volume approximations and Monotonic Upstream-centered Scheme for Conservation Laws (MUSCL) schemes [50]. Though memory requirements can be large for this method, its main source of error comes as a truncation error of the higher order terms in a Taylor expansion of the propagation equations, improving simulation results immensely. Recently, fully kinetic simulations have been used to study near-wall transport, plasma-wave scattering, and plasma turbulence due to kinetic instabilities as potential causes of the anomalous electron transport [31, 41, 51].

- Fluid (Continuum) Approaches: These models solve the moments of the Boltzmann equations and treat each particle species as a fluid continuum, eliminating the statistical noise present in the PIC methods and vastly reducing the computational cost of the direct-kinetic (DK) approach. They are commonly used for equilibrium or near-equilibrium flows. For an

ionized plasma that tracks both collisions and an electromagnetic field, the conservation of mass, momentum, and energy equations can be written as

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\overrightarrow{u}) = S, \tag{1.2a}$$

$$\frac{\partial}{\partial t}(mn\overrightarrow{u}) + \nabla(mn\overrightarrow{u}.\overrightarrow{u} + p) = q(\overrightarrow{E} + \overrightarrow{u} \times \overrightarrow{B}) + \nabla\tau + \overrightarrow{R}, \tag{1.2b}$$

$$\frac{\partial}{\partial t}(n\epsilon) + \nabla \cdot (n\overrightarrow{u}\epsilon + p\overrightarrow{u}) = \nabla \cdot \overrightarrow{Q} + qn\overrightarrow{u} \cdot \overrightarrow{E} + S_{elas} - S_{inelas} + \Phi, \tag{1.2c}$$

where m is the mass, n is the number density, $\overrightarrow{u}$ is the mean velocity, S is the source or sink term, p is the pressure, τ is the viscous stress, $\overrightarrow{R}$ is the collisional friction, ϵ is the mean energy, $\overrightarrow{Q}$ is the conductive heat flux, Φ is the energy dissipation function due to viscous stress, and the subscripts $elas$ and $inelas$ refer to the energy loss due to elastic and inelastic collisions, respectively. The Euler equations treat the right hand side of the above equations as zero while the Navier-Stokes equations retain the viscous terms and heat conduction. In plasma simulations, electromagnetic forces are typically assumed greater than viscous terms while kinetic energy is greater than thermal energy. Further explanation of these equations is provided in Refs. 8 and 50 which discusses the closure problem of fluid equations for terms such as the pressure in the momentum equation. Recent work has determined that kinetic energy cannot be neglected due to a strong ExB drift in Hall effect thrusters.

- Hybrid: In these models, fluid descriptions are often used for particles assuming a Maxwellian energy distribution, often electrons, while PIC methods are used to model the particles in the high-energy tail of the Maxwellian, heavy species, to supplement collisional information. A quasi-neutral assumption is required for the transport equation to calculate the electric field or one must use Poisson's equation to deduce it. These models are highly popular due to their reduced computational cost, seeing such models as Hall2De and HPHall [42, 43, 52]. Other models have seen wave-riding effects where ions become trapped in potential wells, been used to study cathode effects, and azimuthally rotation spokes [53, 54, 55, 56, 32]. The

use of PIC methods can introduce numerical oscillations into the results, but recent studies with hybrid-DK simulations have been performed to mitigate these oscillations [57].

Many of the challenges faced by the plasma community have been realized not just by comparing real-time operational data to ground test data, but also by comparing experimental and computational results. Electron mobility has been discovered as a historically irreconcilable difference between experimental and computational results from erosion rates and spatial density profiles [58, 59]. Facility effects have not been quantified in such a way as to include them in computational models including potential electrical coupling with the facility walls, backsputtering of wall materials, and background pressure effects [60, 61]. Even simply taking reliable measurements of small scale phenomena is not possible with the state-of-the-art modeling techniques [8]. Thus, a truly predictive model continues to elude the community. While one method of combating this quandary is to develop a higher-fidelity model, including any potential numerical noise sources and increased computational costs of such a model, another approach is available.

Using a lower-fidelity model, experimental data can be continually supplied to the model to improve the state of the system. This technique can be used to supplement lower-fidelity models to achieve better quantitative results or with higher fidelity models to determine unknown states or parameters of interest, uncovering uncertain physics. The use of experimental data concurrently with computational models is a process known as data-driven modeling, model-data fusion, or data assimilation.

1.4 Data-Driven Modeling

Data-driven modeling spans all manner of computational practices from state estimation to machine learning and artificial intelligence and can be classified as either offline or online modeling. The supplementation of data can assist classification programs, function approximations, correlation studies, and optimization, amongst others, assuming the existence of considerable and sufficient data describing the system. Artificial neural networks have been used to track film growth during plasma etching processes [62], to estimate sputtered particle distributions for incident ion

energy distributions [63], and to predict deposition yield in plasma spray processes [64]. Sparse regression techniques have been used to develop a functional form of turbulent transport coefficients in plasma processes [65]. Dynamic mode decomposition has become common in the fluids community for the information it provides regarding flow dynamics [66].

While many data-driven methods yield arbitrary, black-box programs that create a relation between inputs and outputs, a subset of data-driven modeling relies on physics-based conservation laws to create more physics-attuned solutions. Physics-informed neural networks have been used to solve supervised learning tasks while respecting physical laws given by nonlinear partial differential equations [67]. More recently, time-embeddings have been used to converge to unknown input parameters for constant or time-averaged quantities in dynamical systems [68]. To understand some of the associated benefits and detractors associated with the methods listed, a discussion of offline and online modeling is presented.

1.4.1 Offline

Offline modeling refers to any method that uses iterative testing to optimize the relation between inputs and outputs or a set of initial conditions for a model. These iterations are performed using sets of test data, or training data, that are used to train the algorithm towards correct relations that can then be applied to datasets of interest. Most common are any form of neural network, mode decomposition, sparsity reducing algorithms, and symbolic techniques. These methods have largely seen applications in image processing and parameter estimation for unknown physical processes. Theory-guided machine learning and physics-informed neural networks have been used to study plasma phenomena as well [69, 67]. More recent work has used low-dimensional representations of physical models to study plasma phenomena either as functional forms [65], mode decomposition [66], or time-embedded phase space plots [68].

The general premise of an offline method is presented in Fig. 1.4. The model is run with a set of values for the unknown input parameters, $\vec{\beta}$, to produce a trial solution, $\vec{x}_{\mathrm{trial}}$. The generated data are compared to the reference solution, $\vec{x}_{\mathrm{ref}}$, and the discrepancy, ϵ, between $\vec{x}_{\mathrm{trial}}$ and $\vec{x}_{\mathrm{ref}}$ is calculated. If the discrepancy, or often called the cost function, is large, a new set of input

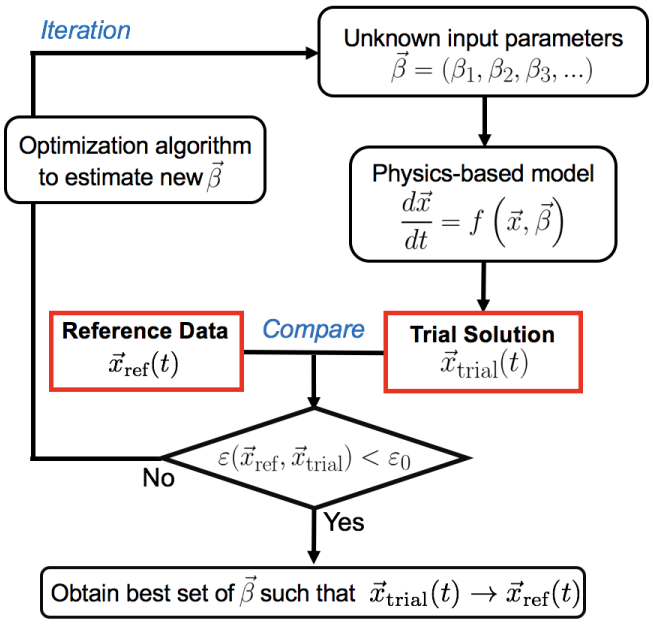

Figure 1.4: An example of an offline data-driven modeling method that relies on the minimization of a cost function.

parameters is determined by the optimization scheme and the calculations are performed again. This iterative cycle continues until a convergence criterion uniquely specified for the given problem is met, i.e., $\epsilon \leq \epsilon_0$, where ϵ_0 is a user-defined convergence criterion. When the minimum cost function is found, it can be considered that an (approximately) optimal set of input parameters which capture the dynamics of the reference solution is obtained.

The three main concerns regarding this data-driven model approach follow directly from Fig. 1.4. The first is what measure should be used to accurately represent the time-dependent system output of the trial and reference solutions ($\vec{x}_{\text{trial}}$ and $\vec{x}_{\text{ref}}$). The second is two-fold: What serves as an appropriate cost function metric between the two discretized solutions (the comparison) and what convergence criterion, ϵ_0, is appropriate. The third is what optimization algorithm best searches the domain and enables complete automation of the optimization process (the iteration). For neural networks these questions become how many nodes and hidden layers should be allowed to exist in the model while for sparse regression and mode decomposition the number of resulting

equations becomes a key question of the model creation.

Work performed prior to this dissertation focused on the development of a phase-space model representation for offline data-driven modeling. A chosen output signal, Fig. 1.5(a), is mapped against itself at an earlier time, Fig. 1.5(b), and subsequently binned as a probability distribution function, Fig. 1.5(c). The Wasserstein Metric, or Earth Mover's Distance (EMD), is used to measure the difference between two probability distribution functions as a single value. Optimization algorithms developed by Sandia National Laboratories [70] are used to iterate through possible parameter sets to find an optimal solution that minimizes the EMD. An interface developed by Stellar Science LLC and the Air Force Research Laboratory called the Galaxy Simulation Builder [71] is used to set the optimization program and coordinate with remote computing systems. This calibration method was shown to recover the reference solution parameters to high levels of accuracy for a variety of nonlinear, dynamic, chaotic computational models. The static, or steady-state, solution generated by this method is useful but limited in its abilities compared to the intricacies of dynamic oscillations in plasma physics.

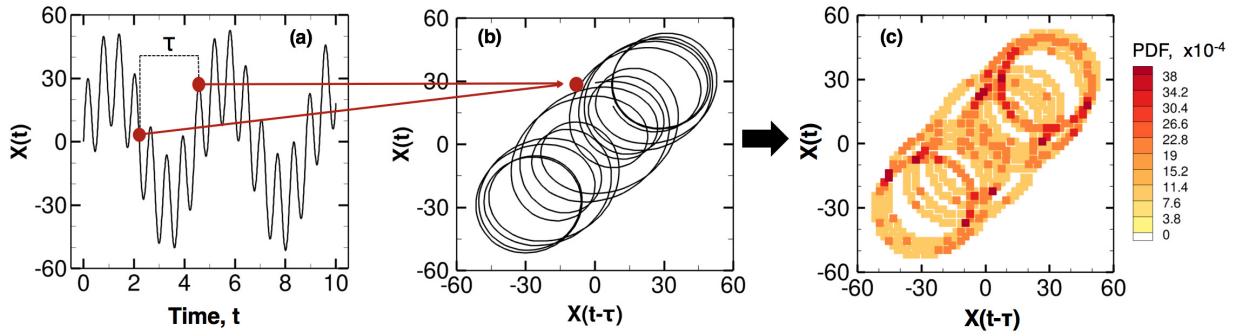

Figure 1.5: An example of the phase-space representation and data comparison model. (a) The time-dependent signal of interest. (b) The phase space representation of the original signal, mapped to two dimensions. (c) The discretized phase space representation to generalize the amount of time the system spends in a region of space.

1.4.2 Online

In comparison, online modeling refers to any estimation technique that operates directly on the system of interest, learning the system relations as it propagates. This approach enables estimation to be performed without knowledge of the entire measurement signal, e.g., the estimation can be performed as measurement data is collected. The most common type of algorithm for this technique is the suite of Kalman filters. These filters range from simple linear filters to more robust particle-based probability filters. Other techniques involve weighted least squares, least median of squares, Bayesian approaches, and generalized maximum-likelihood estimators. The main advantage of any recursive technique is that supplemental data is supplied to the model as it arrives to allow for continuous improvements in the system estimation. Thus, the filter is able to run without knowledge of the entire state of the system at any given time. These estimation techniques often cycle between a prediction and correction phase, using the difference between the current predicted values and newly arrived measurements to correct the system.

The general premise of an online estimation technique is presented in Fig. 1.6. The model is run with the known operating parameters of the system and with some initial guess for any unknown states, $\hat{x}$. As the system propagates forward in time, measurement data, $\tilde{y}$, is supplied continuously or discretely to the model. As a measurement signal arrives, the temporal propagation is paused while the system uses some form of mathematical structure to update the system states from the propagated value (-) to the updated value (+) using the most recent measurement update. Using the improved states, the model returns to the propagation phase, cycling between these two modes until the simulation ends or no further measurement data is available.

Online state estimation is host to its own particular set of challenges. Because they operate in real-time with a physics-based model, these systems are more susceptible to failure. Certain operating conditions often require manual tuning to ensure filter success, increasing the time required to setup the system. Additionally, there are associated computational costs with each estimation technique which are incurred with each run of a simulation compared to a reduced order model being run after training. These algorithms also run the risk of finding an incorrect system state

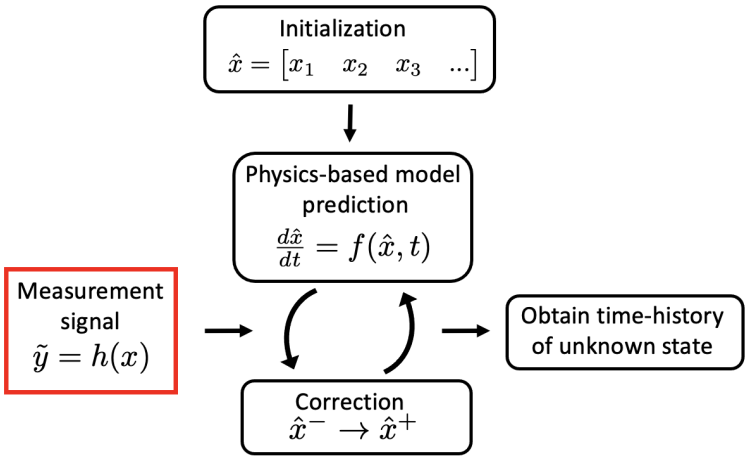

Figure 1.6: An example of an online data-driven modeling method that relies on a predictor-corrector scheme.

based on the known measurements with no inbred method to check the physical reality of the final solution, unlike the cost function of the offline method.

1.5 Problem Statement and Outline

The future of in-space propulsion is dependent on increasing the capabilities of propulsion devices to reach higher thrust levels while also reducing lifespan limiting phenomena. At the required rate of innovation, many of these advances need to be made computationally to prevent tens of thousands of hours of testing and exorbitant facility operation costs. Computational advances are being made to create a truly predictive thruster model that can be accurately used to determine lifespan limiters, thrust profiles, and new design configurations to scale thruster capabilities or investigate phenomena difficult to experimentally measure. Many of these processes in plasmas are related to the high-frequency, short timescale dynamics of electrons. Not only do these processes inform the motion of the electrons as related to thruster erosion, electron mobility, and electron emission from the walls, but also the electron dynamics can affect the overall system behavior. Without experimental data to compare, it is difficult to determine the appropriate high-order physics to include in a model.

Recent advances in computational capability have allowed for the study of data-driven models

which can supplement existing, imperfect models with experimental or on-orbit data to better correlate the computational models to the true physics. The model can be configured to approximate certain processes of interest from general terms that encompass the effects of numerous processes to more targeted applications to study particular dynamics, Previous work performed in the pursuit of a Master of Science degree used an offline technique based on phase space representations of data to optimize static input conditions of a Hall effect thruster model. In this dissertation, a state estimation technique from the suite of Kalman filters that can estimate time-varying states and parameters is developed for plasma physics applications, with the goal of ultimately applying it to Hall effect thruster models.

The remainder of the dissertation is laid out as follows. Chapter II provides a detailed overview of the particular Kalman filter used in this work as well as a discussion of further extensions to the filter. Chapter III presents two simple verification test cases based on known simplistic problems, the Lorentz attractor and the driven-damped harmonic oscillator. These studies are performed to test the filter with sparse measurement signals and chaotic systems. Chapter IV studies a series of global plasma chemistry models that range from simple two-equation xenon models studying low-frequency plasma oscillations to more nonlinear pulsed argon plasma operations. These studies increase the complexity of the propagation model as well as the incoming measurement signals to gain further understanding of the EKF operation given more complex physics. Chapter V extends the Kalman filter into one spatial dimension, studying linear advection and heat conduction models to approximate solving electron mobility coefficients in HETs. In the last section of Chapter V, initial estimation attempts with a one-dimensional fluid HET model are presented to discuss the remaining challenges of this work. Lastly, Chapter VI presents a summary of the presented work and considerations for future research.

2. METHOD

Do let us go on quietly, examining all things and holding fast to that which is good, and trying always to learn more that is useful, and gaining more experience.

- Vincent Van Gogh, *Amsterdam, Apr. 1878*

2.1 Kalman Filters

Kalman filtering has seen wide use since the first Apollo missions where Stanley F. Schmidt is typically credited with famously utilizing the filter as part of the navigation systems [72]. The original filtering method is attributed to Rudolf E. Kálmán and Richard S. Bucy though many other researchers including Stratonovich, Thiele, and Swerling have been understood to have developed similar types of filters [73]. Schmidt is generally credited with the first implementation of the Kalman filter in its traditional two-part application [74].

The suite of Kalman filters are recursive algorithms that move between prediction and correction phases to continually improve estimates of the internal state of dynamic systems using noisy measurements. Both the states and their associated errors are run through this cycle, continually generating *a priori* (before correction) and *a posteriori* (after correction) sets of data. The filter is capable of effectively handling uncertainties in the noisy measurements as well as some uncertainties related to the approximations in the equations that describe the system evolution. The recursive nature of the filter, described later in this chapter, adds to the appeal of the filter by requiring only the most recent estimate to update the new states rather than the entire history of the system states, saving immensely on memory requirements.

The original version of the filter was developed for linear problems as an optimal filter, though since its creation many applications have required the creation of a nonlinear version. The nonlinear extensions of the simple Kalman filter have been used to advance data integrity [75], noise

mitigation [76], and state estimation [77]. These filters have found use in electrical and electronics engineering for x-ray tomography [78], physics for ionosphere modeling and gravitational waves [79], computer science and robotics for movement tracking and sensor fusion [80], industrial engineering for supply chain optimization and fault diagnosis [81], and navigation for GPS autonomy [82].

Perhaps the most useful aspect of a Kalman filter is the freedom allowed with any unknown or uncertain states of interest in the system. Because the filter is expected to use the incoming measurement signals to update the state vector, the states can automatically be assumed to be corrected with every measurement update. Also, because the filter tracks the uncertainty of each state in the system, the propagation equation for any unknown or uncertain physics can be imperfect. In the simplest case, the propagation equation of an unknown state can be considered piecewise constant for all time, i.e., $d/dt = 0$, such that the state is only changed by the discrete measurement updates. This particular propagation may not be suitable for every case of an unknown state in a system, but demonstrates how an imperfect mathematical representation of a state can be used in a Kalman filter due to the measurement correction.

2.1.1 State Space Setup

Kalman filtering is a wide field of stochastic state estimation techniques given a system model and state measurement update. The following derivations are specifically for the continuous-discrete extended Kalman filter where discrete time measurement data are supplied to a continuous propagation scheme.

A general dynamical system can be written using a state vector $\boldsymbol{x}(t)$ as

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), t) + G(t)\boldsymbol{\omega}(t), \tag{2.1}$$

where $\boldsymbol{f}$ is the dynamic model function of the state variables, $G(t)$ is the process noise gain matrix, and $\boldsymbol{\omega}(t)$ is the process noise with covariance $Q(t) = E[\boldsymbol{\omega}\boldsymbol{\omega}^T]$ where $E[]$ is the expectation. The expectation refers to the summation of all possible values of a state, and is considered to represent

the product of probability of an event occurring and the value of the actual observed occurrence. Assuming that the state estimation noise characteristics are described by a Gaussian distribution function, the state estimation is fully described with discrete linear measurements

$$\tilde{\boldsymbol{y}}_k = \boldsymbol{h}\boldsymbol{x}_k + \boldsymbol{\nu}_k, \tag{2.2}$$

where $\tilde{\boldsymbol{y}}$ is the measurement, $\boldsymbol{h}$ is the observation matrix, $\boldsymbol{\nu}$ is the zero-mean measurement noise with covariance $R_k = E[\boldsymbol{\nu}_k \boldsymbol{\nu}_k^T]$, and subscript k denotes the discrete time where the new measurement data are acquired and the Kalman update is performed. The initial state is a random vector with a known mean, denoted by μ_0, and covariance, denoted as $P_0 = E[(\boldsymbol{x}_0 - \mu_0)(\boldsymbol{x}_0 - \mu_0)^T]$.

2.1.2 Extended Kalman Filter

The extended Kalman filter (EKF) is predicated on the following assumptions: randomness can be considered caused by primary random sources exciting states, these primary sources are independent, zero-mean gaussian random processes, and the dynamic systems are linear [73]. Because this is the nonlinear extension to the original Kalman filter, this particular filter the linearization of the system dynamics and measurement equations around the current estimate after each state propagation. For this reason, the EKF is sub-optimal, but has proven effective in practice for systems that are slightly nonlinear.

Figure 2.1 shows a schematic of the continuous-discrete EKF, which is analogous to a predictor-corrector method. A physics-based model is used to continuously propagate the state estimates and associated covariances, while discrete measurement data are used to correct the state and covariance. The predictor step corresponds to temporal evolution of the state vector from $\hat{\boldsymbol{x}}_{k-1}^+$ to $\hat{\boldsymbol{x}}_k^-$ and the corrector step updates $\hat{\boldsymbol{x}}_k^-$ to $\hat{\boldsymbol{x}}_k^+$, where superscript $(-)$ and $(+)$ indicate the estimation before and after the measurement updates, respectively, and subscript k denotes the time step at which measurement data are obtained. Further detail about the derivations of this filter can be found throughout literature [77].

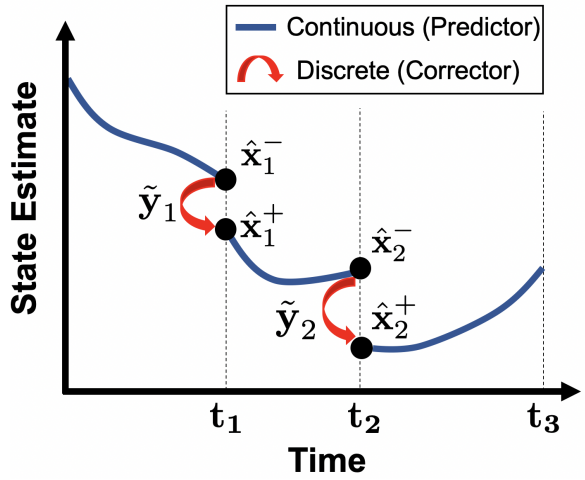

Figure 2.1: A visual representation of the continuous-discrete EKF process over two measurement data points. The model continuously propagates forward until such a time that measurement data arrive and a discontinuous correction is applied to the system.

2.1.2.1 Propagation

Given a discrete stochastic system model and initial conditions, the Kalman filter propagates all states in a system as

$$\dot{\hat{\boldsymbol{x}}}(t) = \boldsymbol{f}(\hat{\boldsymbol{x}}(t), t), \tag{2.3}$$

where f is the nonlinear system model. For the purposes of generating a general propagation, the previous time step, $t = k-1$, and next time step, $t = k$, will be used in the following derivation. This function can be expanded using a Taylor series and written as

$$\boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}) = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}^+) + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}|_{\hat{\boldsymbol{x}}_{k-1}^+}}(\boldsymbol{x}_{k-1} - \hat{\boldsymbol{x}}_{k-1}^+) + H.O.T., \tag{2.4}$$

where the higher order terms (H.O.T.s) are neglected. Note that the partial derivative of f with respect to the states can be written as a Jacobian matrix $F(\hat{\boldsymbol{x}}_{k-1}^+) \equiv \partial \boldsymbol{f}/\partial \boldsymbol{x}|_{\hat{\boldsymbol{x}}_{k-1}}$. Thus, the equation can be rewritten as

$$\boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}^+) = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}^+) + F(\hat{\boldsymbol{x}}_{k-1}^+)\boldsymbol{e_{k-1}},$$

where e_{k-1} contains the residual error of the system, $\boldsymbol{x}_{k-1} - \hat{\boldsymbol{x}}_{k-1}^{+}$.

The expectation of both sides of the previous equation can be taken to yield

$$E[\boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1})] \approx \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}^{+}) + F_{k-1}E[\boldsymbol{e}_{k-1}],$$

where $E[\boldsymbol{e}_{k-1}] = 0$. thus, the time dependent equation for the state vector is simply

$$\dot{\hat{\boldsymbol{x}}}_{k}^{-} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}^{+}). \tag{2.5}$$

The expression created from the Taylor series expansion can be substituted into the forecast error equation, the difference between the true solution and the current estimate as

$$\begin{aligned} \boldsymbol{e}_{k}^{-} &= \boldsymbol{x}_{k} - \hat{\boldsymbol{x}}_{k}^{-} \\ &= \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{\omega}_{k-1} - \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}^{+}) \\ &= F_{k-1}\boldsymbol{e}_{k-1} + \boldsymbol{\omega}_{k-1}, \end{aligned} \tag{2.6}$$

where the representation of $\boldsymbol{x}$ is based upon that of Eq.(2.1). This error equation can be substituted into the error covariance definition as

$$\begin{aligned} P_{k}^{-} &= E[\boldsymbol{e}_{k}^{-}\boldsymbol{e}_{k}^{-T}] \\ &= F_{k-1}E[\boldsymbol{e}_{k-1}\boldsymbol{e}_{k-1}^{T}]F_{k-1}^{T} + E[\boldsymbol{\omega}_{k-1}\boldsymbol{\omega}_{k-1}^{T}] \\ &= F_{k-1}P_{k-1}^{+}F_{k-1}^{T} + Q_{k-1}, \end{aligned} \tag{2.7}$$

to create the propagation equation for the error covariance where $Q(t)$ is the process noise covariance, i.e., the power spectral density of the model uncertainty associated with the physical model. The process noise covariance is a concept that comes from the imperfection of the linearized model with respect to the actual physics. In the absence of a newly arrived measurement data, the filter

continues to propagate, causing covariances to grow.

2.1.2.2 Update

When a measurement becomes available, the filter breaks out of the propagation loop and moves into the correction scheme.

Starting with the measurement equation, the observation function $\boldsymbol{h}$ can be expanded by a Taylor series as

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{h}(\hat{\boldsymbol{x}}^-) + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}|_{\hat{\boldsymbol{x}}^-}}(\boldsymbol{x} - \hat{\boldsymbol{x}}^-) + H.O.T.$$

where the partial derivative can be denoted as $H_k(\hat{\boldsymbol{x}}_k^-) \equiv \partial \boldsymbol{h}/\partial \boldsymbol{x}|_{\hat{\boldsymbol{x}}_k^-}$ and the higher order terms are neglected. The expectation can again be taken on both sides as

$$E[h(\boldsymbol{x})] \approx \boldsymbol{h}(\hat{\boldsymbol{x}}^-) + H_k E[\boldsymbol{e}_k^-], \tag{2.8}$$

where $E[\boldsymbol{e}] = 0$ as before.

If the goal is to approximate the best unbiased estimate, one way is to assume there is some linear combination of both the state and measurement through some unknown form, a, as

$$\hat{\boldsymbol{x}}_k^+ = a + K_k \tilde{y}_k, \tag{2.9}$$

with the unbiased condition of

$$\begin{aligned} 0 &= E[\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^+] \\ &= E[(\boldsymbol{x}_k^- + \boldsymbol{e}_k^-) - (a + K_k \boldsymbol{h}(\boldsymbol{x}_k) + K_k \boldsymbol{\nu}_k)] \\ &= \hat{\boldsymbol{x}}_k^- - a - K_k E[\boldsymbol{h}(\boldsymbol{x}_k)] \\ a &= \hat{\boldsymbol{x}}_k^- - K_k E[\boldsymbol{h}(\boldsymbol{x}_k)], \end{aligned} \tag{2.10}$$

where the term K_k indicates the Kalman gain, a weighting function that will be described later in this section.

This relation can be substituted into the linear combination, Eq. (2.9), to obtain

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + K_k(\tilde{y}_k - E[\boldsymbol{h}(\boldsymbol{x}_k)]), \tag{2.11}$$

which can be combined with the equation for the expectation of h given in Eq. (2.8) to obtain

$$\hat{\boldsymbol{x}}_k^+ \approx \hat{\boldsymbol{x}}_k^- + K_k(\tilde{y}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_k^-)), \tag{2.12}$$

which is the complete update equation for the state vector.

The error of this estimate can be written as

$$\begin{aligned}
\boldsymbol{e}_k &= \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^+ = \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{\omega}_{k-1} - \hat{\boldsymbol{x}}_k^- - K_k(\tilde{y}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_k^-)) \\
&= \boldsymbol{f}(\boldsymbol{x}_{k-1}) - \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1}^+) + \boldsymbol{\omega}_{k-1} - K_k(\boldsymbol{h}(\boldsymbol{x}_k) - \boldsymbol{h}(\hat{\boldsymbol{x}}_k^- + \boldsymbol{\nu}_k) \\
&= F(\hat{\boldsymbol{x}}_{k-1}^+)\boldsymbol{e}_{k-1} + \boldsymbol{\omega}_{k-1} - K_k(H_k(\hat{\boldsymbol{x}}_k^-)\boldsymbol{e}_k^- + \boldsymbol{\nu}_k) \\
&= F(\hat{\boldsymbol{x}}_{k-1}^+)\boldsymbol{e}_{k-1} + \boldsymbol{\omega}_{k-1} - K_k H_k(\hat{\boldsymbol{x}}_k^-)(F(\hat{\boldsymbol{x}}_{k-1}^+)\boldsymbol{e}_{k-1} + \boldsymbol{\omega}_{k-1}) - K_k\boldsymbol{\nu}_k \\
&= (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))F(\hat{\boldsymbol{x}}_{k-1}^+)\boldsymbol{e}_{k-1} + (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))\boldsymbol{\omega}_{k-1} - K_k\boldsymbol{\nu}_k.
\end{aligned} \tag{2.13}$$

This definition of the error estimate can be substituted into the definition of the covariance matrix and simplified as

$$\begin{aligned}
P_k &\equiv E[\boldsymbol{e}_k\boldsymbol{e}_k^T] \\
&= (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))F(\hat{\boldsymbol{x}}_{k-1}^+)P_{k-1}F_{k-1}^T(\hat{\boldsymbol{x}}^+)(I - K_k H_k(\hat{\boldsymbol{x}}_k^-))^T \\
&+ (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))Q_{k-1}(I - K_k H_k(\hat{\boldsymbol{x}}_k^-))^T + K_k R_k K_k^T \\
&= P_k^- - K_k H_k(\hat{\boldsymbol{x}}_k^-)P_k^- - P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T K_k^T + K_k H_k(\hat{\boldsymbol{x}}_k^-)P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T K_k^T + K_k R_k K_k^T.
\end{aligned} \tag{2.14}$$

The optimal value of K_k can be found by minimizing the trace of the covariance matrix with

respect to K such that

$$\begin{aligned} 0 &= \frac{\partial tr(P_k)}{\partial K_k} \\ &= -(H_k(\hat{\boldsymbol{x}}_k^-)P_k^-)^T - P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T + 2K_k H_k(\hat{\boldsymbol{x}}_k^-)P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T + 2K_k R_k, \end{aligned} \tag{2.15}$$

which can be solved for K to obtain

$$K_k = P_k^- H_k^T(\hat{\boldsymbol{x}}_k^-)\left[H_k(\hat{\boldsymbol{x}}_k^-)P_k^- H_k^T(\hat{\boldsymbol{x}}_k^-) + R_k\right]^{-1}, \tag{2.16}$$

which is the well-recognized Kalman gain equation. This gain serves the same purpose as that in a proportional-integral-derivative (PID) controller, directing the filter how to weight the significance of newly arrived measurement data compared to the current prediction. The Kalman gain acts as a weighting factor to correct the states and covariances. It implies that if the *a priori* error covariance is smaller than the measurement noise covariance, the gain will be small and the filter will rely more on the predicted values. Conversely, a larger gain would mean that the measurement noise is smaller than the predicted noise, thus causing the filter to adopt the measurement values.

Substituting this Kalman gain equation back into the posterior covariance equation, Eq. (2.14), yields

$$\begin{aligned} P_k^+ &= (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))P_k^- - (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T K_k^T + K_k R_k K_k^T \\ &= (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))P_k^- - (P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T - K_k H_k(\hat{\boldsymbol{x}}_k^-)P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T - K_k R_k)K_k^T \\ &= (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))P_k^- - [P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T - K_k(H_k(\hat{\boldsymbol{x}}_k^-)P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T + R_k)]K_k^T \\ &= (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))P_k^- - [P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T - P_k^- H_k(\hat{\boldsymbol{x}}_k^-)^T]K_k^T \\ &= (I - K_k H_k(\hat{\boldsymbol{x}}_k^-))P_k^-, \end{aligned} \tag{2.17}$$

which is the final form of the update equation for the covariance matrix. Note that the measurement does not have to directly correlate to the states. Instead, the measurement may be some combina-

tion of states or simply an equation involving a single state, captured in the basis function H_k. The filter returns to the propagation phase until a new measurement signal becomes available.

2.2 Physics-Informed Constraints

State and parameter estimation necessitates robust and efficient calculations that take physical and mathematical constraints into account. For instance, the time intervals of measurement acquisition may not be equal, as indicated in Figure 2.1, or the estimated state and variance may become unphysical without any constraints on the filter.

In the EKF, two key *ad hoc* parameters are the process noise covariance, Q_{k-1}, which is applied to the system in the continuous covariance propagation equation (Eq. (2.7)), and the measurement noise covariance, R_k, which is used in the discrete measurement update (Eq. (2.17)). These parameters are instrumental in the fidelity of the estimation as well as the robustness of the filter against nonlinearities or sparse measurements. While R_k is usually quantified from measurement instrument uncertainties, the core ability of the EKF relies on the choice of Q_{k-1}, which describes the imperfection of the *model* with respect to the *physics*.

Furthermore, filter divergence, where the solution tends far from the measured quantities or the filter never completes, can be caused by a multitude of uncertainties in both the filter and the nonlinear model [83]. The nature of the linearization used in this filter typically requires the initial guess of the estimator to be close to the true solution to initially prevent the filter from diverging [77]. Then, selected initial covariances must be large enough to allow the system to provide a state estimation within the confidence bound. Note that for the remainder of this chapter, the subscripts for R_k and Q_{k-1} will be dropped when discussing the variable within the text as alternate subscripts may be used in equations for clarity.

2.2.1 Design of the Process Noise Covariance

The extended Kalman filter is almost entirely reliant on two ad hoc parameters. The first of these is the process noise covariance, Q, which quantifies the uncertainties in the physics-based model, whether from an incomplete model or numerical uncertainties. The chosen Q can specif-

ically be seen as the higher-order terms that are truncated when selecting the nonlinear dynamical model, i.e., the difference between the true nonlinear dynamics and the model, as shown in Eq. (2.1). Some of the most common practices for designing this matrix involve guess-and-check methods, adaptive solutions, neural network training, and continuous white noise approximation. Many studies have performed manual searches for a process noise covariance value that returns the desired performance from the filter [84] by applying machine learning to find optimal values of the noise covariances [63, 64] or by using an adaptive Kalman filter [85, 86, 87]. These methods not only require additional computational cost and involve complex models, but also fail to impose any constraints on the filter to ensure robust performance. Quadrature methods have been used to minimize the effects of linearization error, leading to solutions such as the unscented Kalman filter. Other studies have attempted to constrain the estimates of the filter based on equality or inequality constraints known by the user for a given system [88, 89, 90]. Adaptive techniques that use dynamic noise covariances (particularly Q) are a common approach, e.g., scaling factors [91].

This work proposes the imposition of physical and mathematical constraints on the process noise covariance. This covariance value is allowed to vary in time to more realistically account for time-dependent, fluctuating noise sources. The filter is thus designed to ensure that the variances P_{mm} ($m \in \mathcal{N}$) are non-zero, positive values while also adding a constraint to prevent unwanted solutions such as negative values of states, including density and temperature.

The process noise covariance is related to the uncertainty associated with the *imperfection* of the physics-based model. Equation (2.7) is where the linearization of nonlinear dynamics is taken into account, using the Jacobian matrix of the nonlinear physics-based model. This suggests that the process noise covariance is directly correlated with the truncated higher-order terms of the nonlinear dynamics. For simplicity, the relationship between the process noise variance and the state variance can be written as

$$(P_{mm})_{k+1}^{-} - (P_{mm})_{k}^{+} = Q_{k \to k+1} \triangle t, \tag{2.18}$$

where m denotes the unknown variance element, the superscripts denote the predicted $(-)$ and

corrected (+) values, and $Q_{k\to k+1}$ and $\triangle t = t_{k+1} - t_k$ are the process noise covariance and the time step between two consecutive measured data, respectively (note that $Q_{k\to k+1}$ is assumed to be constant between t_k and t_{k+1}). The choice of Q is critical as the system uncertainty must increase when the data acquisition is infrequent since the propagated uncertainties are likely to be accumulated over time. However, a large variance could result in a negative solution for one or more state estimates, potentially violating some physical constraints, e.g., density, temperature, and pressure must be positive at all times. In the proposed EKF model, Q is monitored to satisfy the mathematical constraint, i.e., the variances must be non-negative, and prevent divergence of the filter estimation. This is done by setting the unknown state covariance as constant at all times t_{k+1} and solving Eq. (2.18) for Q. Further information regarding the constraints placed on these noise covariances are provided in the following section in the context of the first testcase.

When the original EKF with a fixed *ad hoc* signal process covariance is used, the uncertainties at the unknown state, P^-_{mm}, are dependent on the time interval, as discussed in Eq. (2.18). On one hand, it is correct that the covariance, i.e., uncertainty, must increase and the Kalman gain consequently approaches unity for the measured state, $K_{ll} \to 1$, as the time interval of measurement update increases. Since the estimation from the model accumulates uncertainties, the experimental data are trusted more than the predictions of the physics-based model. However, a large system covariance means that the state variable could be negative during the continuous prediction phase, which will violate the physical constraints depending on the problem of interest. Alternatively, if the system variance approaches zero, the Kalman gain approaches zero, i.e., $K_{ll} = (R/P_{ll} + 1)^{-1} \to 0$, accounting for a finite measurement noise. This approach to zero indicates that the EKF trusts the prediction over the measurement. But, a frequent data acquisition should not mean that the model prediction is more correct than the measurement. Since the definition of the signal process noise is due to the *imperfection* of the model with respect to the *physical* data, i.e., $G(t)\boldsymbol{\omega}(t) = \dot{\boldsymbol{x}}(t) - \boldsymbol{f}(\boldsymbol{x}, t)$, Q may be affected by the noisy measurement. Hence, even if the time integral of data acquisition is small, it could be more physical to account for a finite system uncertainty before the measurement update. In other words, the uncertainty of the

prediction may not be dependent on how frequent the Kalman update is performed and therefore should have some independency from the size of the measurement time step. By addressing this concern, the PC-EKF is therefore a robust way to incorporate physics for both smaller and larger data acquisition time intervals.

2.2.2 Design of the Measurement Noise Covariance

The secondary *ad hoc* parameter is the measurement noise covariance, R. The physical phenomena are experimentally measured using sensors or other measurement diagnostics equipment that have some inherent uncertainties, typically supplied by the manufacturer. As with the process noise covariance, ideally the errors assumed by manufacturers would be suitable values for this covariance. This may not always be the case, or the solutions may perform better is allowed to adapt and change in time.

Using the same requirements of non-zero, positive covariance values, the measurement noise covariance can be updated in two manners.

For any given measurement, the uncertainties associated with data acquisition must be evaluated, i.e., $\boldsymbol{\nu}(t)$ in Eq. (2.2). As shown in Eqs. (2.12) and (2.17), the Kalman gain is reduced in the presence of a large measurement noise, leading to adopting the predicted estimate, $\hat{\boldsymbol{x}}^{-}$, compared to the measurement, $\tilde{\boldsymbol{y}}_k$. Statistical uncertainties of an instrument are defined by a noise distribution that is defined for a set of measurement conditions. In the EKF approach, such noise is assumed to be Gaussian with a certain amplitude.

One can evaluate the constraints associated with the time-dependent measurement noise, σ_R, applied to the measurement signal, x_1, through the R_{11}-component in this testcase as $R_{11} = \sigma_R^2$. The Kalman gain for the case where the measurement is the first element in the state vector $\hat{\boldsymbol{x}}$, i.e. $H_k = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ for all time, k, is given by

$$K_k = \begin{bmatrix} P_{11}^- & P_{12}^- & P_{13}^- \\ P_{21}^- & P_{22}^- & P_{23}^- \\ P_{31}^- & P_{32}^- & P_{33}^- \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \left[R_{11} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} P_{11}^- & P_{12}^- & P_{13}^- \\ P_{21}^- & P_{22}^- & P_{23}^- \\ P_{31}^- & P_{32}^- & P_{33}^- \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right]^{-1} = \begin{bmatrix} P_{11}^- \\ P_{21}^- \\ P_{31}^- \end{bmatrix} \frac{1}{R_{11} + P_{11}^-}. \tag{2.19}$$

The covariance update equation, in Eq. (2.17), can be written as

$$P_k^+ = \left[\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \right] \begin{bmatrix} P_{11}^- & P_{12}^- & P_{13}^- \\ P_{21}^- & P_{22}^- & P_{23}^- \\ P_{31}^- & P_{32}^- & P_{33}^-. \end{bmatrix}. \tag{2.20}$$

The variance P_{33}^+ term, using the Kalman gain from Eq. (3.4), can be written as

$$P_{33}^+ = P_{33}^- - \frac{P_{13}^- P_{31}^-}{R_{11} + P_{11}^-}. \tag{2.21}$$

Note that $P_{ij} = P_{ji}$, where $i \neq j$, due to the symmetry in the system covariance matrix. It is immediately evident that the variance corrected by the measurement will always be less than the predicted variance because the second term on the right-hand side in Eq. (2.21) is positive. In other words, the measurement is used to reduce the uncertainty of the estimate, i.e. $P_k^+ < P_k^-$. However, the standard deviation, and thus the variance, of each state, i.e. P_{ii}, must be a positive value at *all* times. Therefore, $P_{ii}^+ > 0$ is a condition that must be satisfied.

In addition, without any constraints on R_{11}, there could be a case where P_{33}^+ becomes negative as can be seen from Eq. (2.21), and cause the failure of the EKF. Noting that the measurement noise itself may be uncertain, one can satisfy a positive system variance by adjusting the Kalman update. For instance, if R_{11} is infinitely large, namely, the measurement uncertainty is large, $P_{33}^+ = P_{33}^-$, the measurement is not trusted and the prediction from the physics-based model is adopted. On the other hand, if the measurement error is small, i.e. R_{11} is small, there could be a situation where P_{33}^+ becomes negative, which can make the EKF unstable. To keep P_{33}^+ positive, the measurement

noise, R_{11}, is designed to be adjusted (when needed) as follows:

$$R_{11} \geq \frac{\left(P_{13}^{-}\right)^2}{P_{33}^{+} - P_{\min}} - P_{11}^{-}, \tag{2.22}$$

where the $P_{\min}$ is a user-defined value for the minimum allowed unknown state covariance which increases R_{11} to keep $P_{33}^{+} > 0$. Note that this minimum variance value is imposed on the variance of the unknown state, P_{mm}, where m is the index corresponding to the unknown state, i.e. $\min(P_{mm}^{+}) = P_{\min}$.

Alternatively, the measurement noise covariance can be assumed as some percentage of the corresponding state, calculated at each measurement update phase.

2.3 Extensions of the Kalman Filter

Further extensions of the Kalman filter exist to better handle nonlinear systems. Within the EKF framework, the filter can be applied in the modal space, with a finite element approach, or with a finite difference approach. Beyond the EKF exists the iterated extended Kalman filter (IEKF) which acts exactly as the EKF, except that it iterates through multiple state estimates to ensure an optimal update is reached. This is performed by tracking the difference between update values, stopping the iteration once the solution no longer changes. This filter has been shown to act identically to the EKF unless further modifications are made, such as using the maximum a posteriori (MAP) criterion or a quasi-Newton technique. There are also unscented Kalman filters (UKF) which uses a deterministic sampling technique known as the unscented transformation (UT) to pick a minimal set of sample points (called sigma points) around the mean that are then propagated through the equations to determine a new mean and covariance estimate. While more robust against nonlinearities, this filter can face difficulties with diverging sigma points as well as increased computational cost. The ensemble Kalman filter (EnKF) replaces the covariance matrix with a sample covariance. The EnKF is a particular type of particle filter, assuming that all probability distribution functions are Gaussian. Ensemble Kalman filters represent the distribution of the system state using a collection of state vectors, called an ensemble, and replace the covariance

matrix by the sample covariance computed from the ensemble. The ensemble is operated with as if it were a random sample, but the ensemble members are really not independent. Each extension has its merits, but as the EKF appears suitable for this research, the added computational time of a more robust filter was found unnecessary.

3. ZERO-DIMENSIONAL VERIFICATION

I now consider myself to be at the beginning of the beginning of making something serious.

- Vincent Van Gogh, *Etten, Dec. 1881*

The study of zero-dimensional models is a broad topic of research as nearly every physical system can be reduced to a simple, time-dependent model. For smooth dynamical systems, these models often appear as attractors or predator-prey models. Computational research communities have long used these simplified models as initial test cases for proving new methods or modeling approaches due to the commonality of these models as well as the wealth of information that can be learned from them. For some global models, studies of phase shifts and oscillation amplitudes can provide significant understanding of how well an applied system, such as a data-driven model, can recreate the results. Other models, such as chaotic attractors, are often used to demonstrate the ability to filter noise or to track system states without learning some assumed, repeated pattern. This chapter presents three simple test cases to demonstrate the capabilities of the physics-constrained EKF presented in the previous chapter. These test cases are used to benchmark against an original EKF where the ad hoc noise covariances remain constant, study the effects of sparse measurement updates, and demonstrate the abilities of the constrained EKF to estimate unknown states in plasma physics applications.

3.1 Lorenz System

The Lorenz system was first developed in the early 1960s and is attributed to work performed by Edward Lorenz. The original intention for the system was to develop a simplified mathematical model for atmospheric convection to study meteorological events. In particular, the Lorenz system models a two-dimensional layer of fluid that is uniformly warmed from below and heated from

Part of this chapter is reprinted with permission from "Real-time state estimation of low-frequency plasma oscillations in Hall effect thruster" by C.M. Greve, M. Majji, and K. Hara, 2021, *Physics of Plasmas*, 28, 093509, COPYRIGHT 2021 by AIP Publishing.

above by tracking the rate of convection and temperature variation in both horizontal and vertical directions. The system can be shown to never cross the same point in space more than once (deterministic) as it orbits around one or more attractor points at fixed locations in space. the system is non-periodic because the system will switch attractors in an unpredictable manner. The intrigue of this system is that use of the same three constant parameters with imperceptibly different input parameters will yield different solution trajectories as time advances due to the chaotic nature of the system. Thus, this system is unpredictable over time although its general dynamics can be assumed based on the known attractor points of the system. Since its development, these equations have seen similar application in fields of lasers [92], DC motors [93], chemical reactions [94], and the Malkus waterwheel [95].

The Lorenz system is used in this dissertation to study the convergence properties of the physics-constrained EKF (PC-EKF) in comparison with the original EKF where fixed *ad hoc* noise covariances are used. To demonstrate the capabilities of the EKF models, a self-verification test case is created where a time-dependent solution obtained from a particular calculation of the physics-based model is taken as a *reference* solution, which is provided in the EKF models as the measurement data, $\tilde{\mathbf{y}}$. Quantitative assessments of the filters are made by comparing the reference solution and the calculated solution obtained from the EKF models before applying them to plasma dynamical models. Two tests conducted in this section include: (i) to test how the EKF models handle added noise (random fluctuation) in the measurement for the state and parameter estimation and (ii) to test how the EKF models handle sparse measurement dataset, i.e., when the Kalman update is conducted at a lower frequency. These capabilities play an important role in application of the filters to physical systems with measurement uncertainties and limited data acquisition compared to the dynamics of interest.

3.1.1 The Lorenz Equations

The Lorenz system serves as a common test case for chaotic nonlinear system studies due to its non-periodic yet deterministic nature. Extensive literature can be found on the derivation of the system, its applications, and testing, so only a brief discussion of the equations is provided [68, 96,

97, 98]. The Lorenz equations are given as

$$\frac{dX}{dt} = \sigma_L(Y - X), \tag{3.1a}$$

$$\frac{dY}{dt} = X(\rho_L - Z) - Y, \tag{3.1b}$$

$$\frac{dZ}{dt} = XY - \beta_L Z, \tag{3.1c}$$

where, for this work, X, Y, and Z are spatial positions of the system that vary in time, t, and σ_L, ρ_L, and β_L are constant parameters that govern the overall shape and location of the attractor lobes. The values $\sigma_L = 10$, $\rho_L = 28$, and $\beta_L = 8/3$ are used across literature to produce the well-recognized butterfly attractor solution of the Lorenz equations [68] and will be used in this study for consistency.

3.1.2 Simulation Setup

Figure 3.1 shows a comparison of (a) a smooth Lorenz system without noise, (b) a dataset sampled using a time interval of 0.01 with added noise, and (c) a sparse dataset using a time interval of 0.1 with added noise. The latter two cases are used in this study to estimate $\hat{\sigma}_L$ using both the original EKF and PC-EKF models. A low frequency data acquisition can result in the lack of defined attractor lobes or any coherent structure to the Lorenz system Because the shape of the attractor is specific to the unknown parameter of interest, this loss of definition can be detrimental to the estimation of the dynamical system, particularly with sparse data.

In Ref. 68, the development of an inverse problem to obtain all three parameters was demonstrated using an offline estimation technique based on a first Wasserstein metric cost function. Recent work by Dubois et. al. uses a recurrent neural network to predict solutions multiple time steps ahead of the current state without requiring the calculation of error statistics such as in Kalman filters [99]. Sparse regression techniques have also been used to estimate the states of the Lorenz system [100].

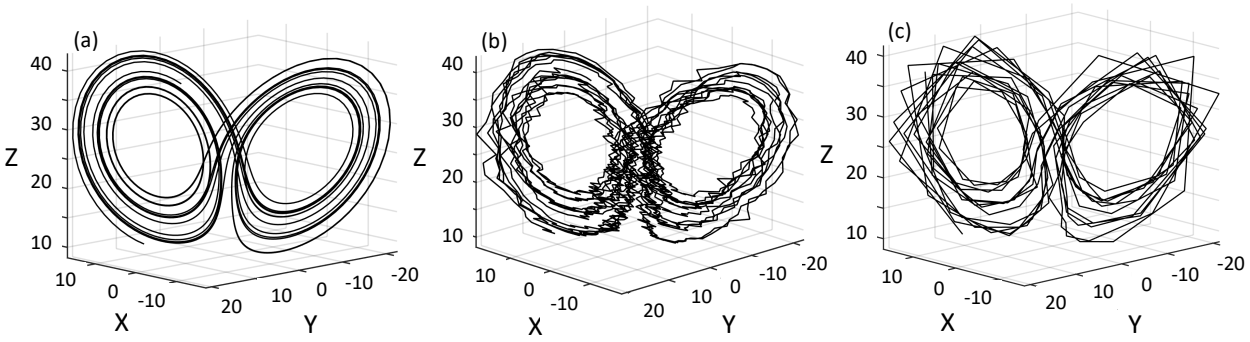

Figure 3.1: A comparison of Lorenz solutions for (a) the original data without noise, (b) 100 samples per unit time measurement data with noise, and (c) 10 samples per unit time measurement data with noise. The random noise has a maximum value of 1.

The present study is performed to investigate whether the physics-constrained EKF can estimate the input parameter σ_L, given different data acquisition frequencies. The input parameter σ_L is included as a component in the state variable vector of the EKF such that $\boldsymbol{x} = \begin{bmatrix} X & Y & Z & \sigma_L \end{bmatrix}^T$. This state vector leads to the following Jacobian matrix

$$F(\boldsymbol{x}(t), t) = \begin{bmatrix} -\sigma_L & \sigma_L & 0 & (Y - X) \\ \rho_L - Z & -1 & -X & 0 \\ Y & X & -\beta_L & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.2}$$

where the unknown state σ_L is considered piecewise constant during the continuous propagation phase.

The X state calculated using a complete Lorenz model with $\sigma_L = 10$ and added white noise is used as the measurement signal, $\tilde{\boldsymbol{y}}_k$, to estimate the state variables. In both EKF models, σ_L is taken into account within the state vector so its uncertainty, i.e., variance, can be propagated in time. For this study, we assume that we do not know the underlying dynamics associated with the unknown parameter, σ_L. Thus, σ_L is assumed piecewise constant, i.e., $d\sigma_L/dt = 0$, during the continuous propagation phase of the EKF. This condition means that the σ_L estimate is only

updated based on the newly arrived measurement data during the Kalman update.

This test case creates a reference solution using the coefficients $(\sigma_L, \rho_L, \beta_L) = (10, 28, 8/3)$ with initial condition $(X, Y, Z) = (10, 10, 10)$. The goal of this test case is to estimate the value of the σ_L parameter using the X state with added noise as the measurement data. This test case is chosen to demonstrate the ability of the solver to estimate an input parameter of the system that cannot be observed by experimental data. Hence, it serves as a self-verification test. Artificial random fluctuations with maximum amplitude of 1 are added to the reference solution of X, i.e., $\tilde{\boldsymbol{y}} = X_{\text{ref}} + \nu$ where ν is a random number between 0 and 1, to test the whether the filter handles to handle noisy data.

Both the reference solution and the estimation are run for 20 dimensionless units of time using a fourth-order Runge-Kutta solver and time step of 0.01 unless otherwise stated. Because the σ_L parameter is assumed constant during the continuous prediction phase, all of the σ_L elements in the Jacobian matrix are treated as zero in the model prediction equations while the derivatives of the other three state are taken into account. The estimate initial conditions are set using values that are not identical to the reference solution, namely, $(\hat{X}, \hat{Y}, \hat{Z}, \hat{\sigma}_L) = (15, 3, 6, 20)$, while the diagonal variance values are initialized to $P_{mm} = 10$, with the non-diagonal components equal to zero.

3.1.2.1 Physics-Constrained EKF Setup

Based on the discussion presented in the previous chapter, the variance of the unknown state, $P^-_{mm,k+1}$, where $m = 4$ corresponds to the desired unobserved input parameter σ_L, is set as a constant, P_0, so that Q, the model error covariance, can vary in time. This is equivalent to adapting the Q value at every prediction stage between two consecutive measurement updates, as can be seen from Eq. (2.18). As discussed in the previous chapter, setting P_0 only affects the final propagation time step covariance value before new measurement data arrive, i.e., P^-_k. This particular variance term will be updated by the Kalman update and allowed to grow during the continuous propagation. The constant P_0 value is set to give an upper limit to how much the uncertainty is allowed to grow.

One can evaluate the constraints associated with the time-dependent measurement noise, σ_R, applied to the measurement signal, X, as $R = \sigma_R^2$. As the measured state for this test case directly

relates to the first component of the state vector, the corresponding basis function can be written as $H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$. Solving Eqs. (2.16) and (2.17) leads to the following relation for the unknown state covariance:

$$P_{44}^{+} = P_{44}^{-} - \frac{P_{14}^{-} P_{41}^{-}}{R + P_{11}^{-}}. \tag{3.3}$$

It is immediately evident that the variance corrected by the measurement will always be less than the predicted variance because the second term on the right-hand side in Eq. (3.3) is positive due to the symmetry of the covariance matrix and a positive-definite variance. In other words, the measurement is used to reduce the uncertainty of the estimate, i.e., $P_k^{+} < P_k^{-}$. However, the standard deviation, and thus the variance, of each state, i.e., P_{mm}, must be a positive value at *all* times. Therefore, $P_{mm}^{+} > 0$ is a condition that must be satisfied.

Noting that the measurement noise itself, σ_R, may be uncertain, one can satisfy a positive system variance by adjusting the Kalman update. For instance, if R is infinitely large, the measurement is not trusted and the prediction from the physics-based model is adopted, causing little, if any, change to the current state and covariance estimates. On the other hand, if the measurement error is small, i.e., R is small, any of the state variances, P_{mm}, could be *over-corrected* and become negative as can be seen from Eq. (3.3). This potential flip in variance sign is not only ill-posed, but indicates that the EKF is unstable. To keep these diagonal terms positive, especially those related to the unknown states, P_{44}^{+} in this test case, the measurement noise covariance, R, is designed to be adjusted at any point in time as follows.

The Kalman gain for the case where the measurement is the first element in the state vector $\hat{\boldsymbol{x}}$, i.e. $H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$ for all time, k, is given by

$$K_k = \begin{bmatrix} P_{11}^- & P_{12}^- & P_{13}^- & P_{14}^- \\ P_{21}^- & P_{22}^- & P_{23}^- & P_{24}^- \\ P_{31}^- & P_{32}^- & P_{33}^- & P_{34}^- \\ P_{41}^- & P_{42}^- & P_{43}^- & P_{44}^- \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \left[R_k + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} P_{11}^- & P_{12}^- & P_{13}^- & P_{14}^- \\ P_{21}^- & P_{22}^- & P_{23}^- & P_{24}^- \\ P_{31}^- & P_{32}^- & P_{33}^- & P_{34}^- \\ P_{41}^- & P_{42}^- & P_{43}^- & P_{44}^- \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right]^{-1} = \begin{bmatrix} P_{11}^- \\ P_{21}^- \\ P_{31}^- \\ P_{14}^- \end{bmatrix} \frac{1}{R_k + P_{11}^-}. \tag{3.4}$$

The covariance update equation, in Eq. (2.17), can be written as

$$P_k^+ = \left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} K_{k,1} \\ K_{k,2} \\ K_{k,3} \\ K_{k,4} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \right] \begin{bmatrix} P_{11}^- & P_{12}^- & P_{13}^- & P_{14}^- \\ P_{21}^- & P_{22}^- & P_{23}^- & P_{24}^- \\ P_{31}^- & P_{32}^- & P_{33}^- & P_{34}^- \\ P_{41}^- & P_{42}^- & P_{43}^- & P_{44}^-. \end{bmatrix}. \tag{3.5}$$

The variance P_{44}^+ term, using the Kalman gain from Eq. (3.4), can be written as

$$P_{44}^+ = P_{44}^- - \frac{P_{14}^- P_{41}^-}{R_k + P_{11}^-}. \tag{3.6}$$

Note that $P_{ij} = P_{ji}$, where $i \neq j$, due to the symmetry in the system covariance matrix. It is immediately evident that the variance corrected by the measurement will always be less than the predicted variance because the second term on the right-hand side in Eq. (3.6) is positive. In other words, the measurement is used to reduce the uncertainty of the estimate, i.e. $P^+ < P^-$. However, the standard deviation, and thus the variance, of each state, i.e. P_{ii}, must be a positive value at *all* times. Therefore, $P_{ii}^+ > 0$ is a condition that must be satisfied. This equation can be rewritten as

$$R \geq \frac{\left(P_{14}^-\right)^2}{P_{44}^+ - P_{\min}} - P_{11}^-, \tag{3.7}$$

where the $P_{\min}$ is a user-defined value for the minimum allowed unknown state covariance. Setting a minimum, positive, value for this unknown variance increases the measurement noise uncertainty

to ensure $P_{44}^{+} > 0$. Note that this minimum variance value is imposed on the variance of the unknown state, P_{mm}, where m is the index corresponding to the unknown state, i.e., $\min(P_{mm}^{+}) = P_{\min}$. As this is a value to ensure positivity, P_{min} can simply be set as some desired value based on general intuition, removing the necessity to tune an *ad hoc* parameter while creating a more robust estimation setup.

As both the original EKF and the physics-constrained extended Kalman filter (PC-EKF) will be used for this test case, the initial measurement noise covariance, R, is set to 3 and the initial process noise covariance, Q, is set to 15. For the physics-constrained EKF, $P_0 = 10$ for P_{44}^{-} and the minimum covariance value introduced in Eq. (3.7) is set to $P_{\min} = 0.1$.

3.1.3 Results

Figure 3.2 shows the temporal evolution of the PC-EKF state estimates $(\hat{X}, \hat{Y}, \hat{Z})$ with their corresponding uncertainties, calculated from the variances from the system covariance matrix, compared to the reference solutions $(X_{\text{ref}}, Y_{\text{ref}}, Z_{\text{ref}})$ using a measurement frequency of 10 samples per unit time, shown in Fig. 3.1(c). Note that the dark red line is the estimated solution of interest while the pink shade denotes the uncertainty bounds. The reference solutions are overlaid in the black dashed line. The small variance bounds indicate a high level of confidence in the filter dynamics and can be seen to remain small throughout most of the estimation. Looking at Fig. 3.2(a), the short transient time between the larger initial variance values to the smaller adaptive variances is clearly visible. This reduction in variance is indicative of the filter finding the measurement dynamics. This short transient time despite sparse, noisy measurement data further evidences the robustness of the PC-EKF for estimating system dynamics. Note that the slight ballooning of the uncertainty between the time units 2 and 3 corresponds to a poorer estimation of the measured state during that time, i.e., $\hat{\boldsymbol{x}}_X$ is farther from the reference solution than at other times during the simulation.

Figure 3.3 shows the PC-EKF estimation of the input parameter σ_L, as it is updated over time for two different cases where the measurement data arrive frequently at 100 samples per time unit (as shown in Fig. 3.1(b)) and where the measurement data are sparse at only 10 samples per

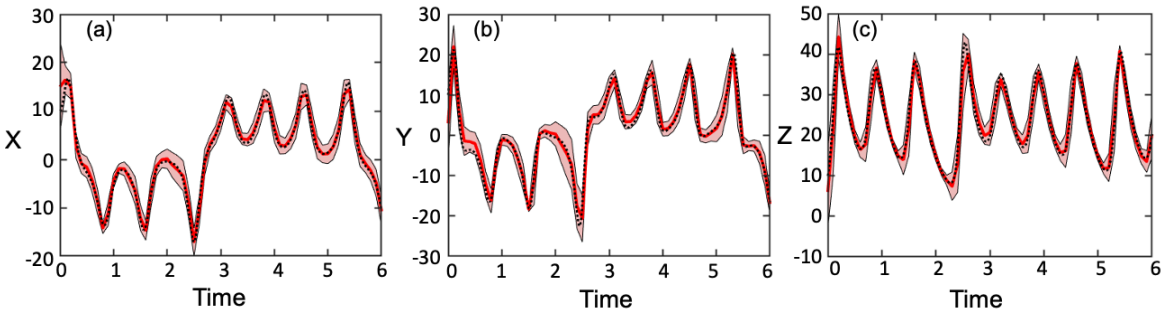

Figure 3.2: The estimation results and reference solutions for the Lorenz system state variables (a) X, (b) Y, and (c) Z run using a measurement frequency of 10 samples per unit time, corresponding to the dataset shown in Fig. 3.1(c). The reference solution X_{ref}, Y_{ref}, and Z_{ref} (black dash) is overlaid with the estimate $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ (red) and corresponding uncertainty bounds (pink).

time unit (as shown in Fig. 3.1(c)). The estimate of $\hat{\sigma}_L$ fluctuates in time due to the continued measurement updates performed by the filter and is shown to never settle into a true, steady-state condition. This is due to the assumed measurement noise and system noise used by the Kalman filter updated compared to the true amount of noise in the measurement. After the initial transient period during the first 1-2 time units, the solution oscillates near the true value of the estimate and the true reference solution is always within the uncertainty bounds. The time-averaged $\hat{\sigma}_L$ approaches the correct $\sigma_{L,\text{ref}} = 10$ with the high frequency measurement case averaging to 9.4648 and the low frequency case averaging to 9.2385. It is to be noted that due to the chaotic nature of the Lorenz system, the time-averaged $\hat{\sigma}_L$ varies dependent on the initial condition, particularly when the measurement data are sparse. This led the original EKF to fail, which is discussed shortly.

Figure 3.4 shows the effects of data acquisition on the performance and robustness of both the original EKF and the physics-constrained EKF models by comparing the root mean square (RMS) errors of the unknown state, $\hat{\sigma}_L$. As discussed in Fig. 3.3, the time-averaged $\hat{\sigma}_L$ can vary with sparse data based on a different measurement acquisition frequency, the initial condition, and the measurement noise due to the chaotic nature of the Lorenz system. The results presented in this figure use identical initial conditions with randomly applied, normally distributed noise in the exact state, X, to create the measurement. Due to this random noise, no two runs are identical despite identical initial conditions. For this reason, multiple tests are run to remove the variation in solution

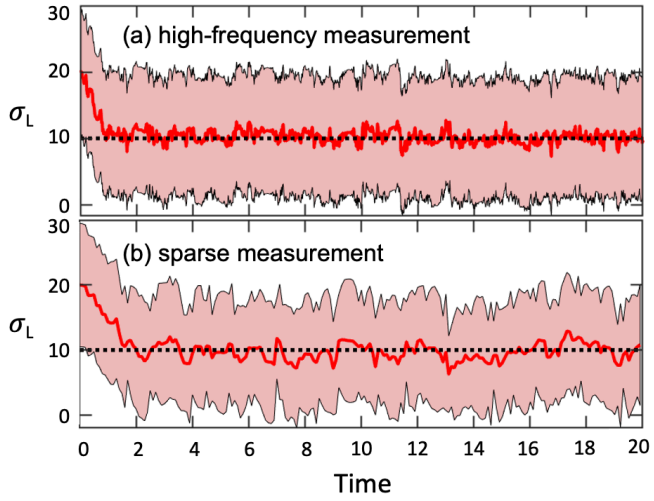

Figure 3.3: The σ_L parameter estimation results for (a) 100 samples per unit time, corresponding to Fig. 3.1(b), and (b) 10 samples per unit time data acquisition frequencies, corresponding to Fig. 3.1(c). The reference solution (black dash) is overlaid with the estimate (red) and the uncertainty bounds (pink).

caused by the sensitivity of the system to the incoming measurement [68]. The presented RMS error of the time-averaged $\hat{\sigma}_L$ reflects the average of 15 individual tests at each data acquisition frequency. The second half of the estimation, i.e., $t \in [10, 20]$, is used to calculate these averages to remove the effects of the initial transient period.

It can be seen from Fig. 3.4 that both the original EKF and PC-EKF perform well when the measurement data stream is fully resolved in time. The original EKF is shown to produce lower RMS error values for the majority of successful runs compared to the PC-EKF, demonstrating the abilities of this filter. However, it is notable that the original EKF becomes unstable at lower data acquisition frequency, i.e., with sparse data. As the time steps between consecutive measurement updates become larger, the uncertainties during the continuous prediction phase become unphysical, causing the propagation phase to fail. Once the simulation fails for as many as one out of every five runs, the results are no longer plotted, which is the case for the original EKF model results. These failures occur when the estimate reaches an unphysical value and causes the EKF to diverge. It is important to note that the choices of R and Q in the original EKF model are not altered between cases, thus becoming non-optimal and largely affecting the performance of the

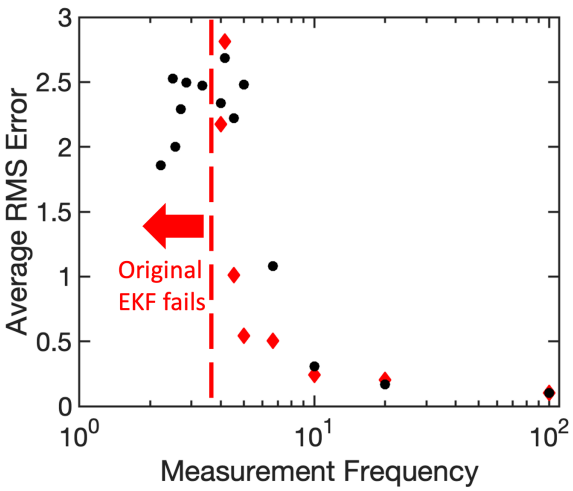

Figure 3.4: The root-mean-square error of the time-averaged $\hat{\sigma}_L$ value obtained from 15 separate runs for the PC-EKF (black circle) and original EKF (red diamond). The EKF solutions are not shown below a measurement frequency of 5 since a finite number of runs failed while the PC-EKF exhibits robust state estimation with sparse data.

filter [84, 85]. Were these values to be varied, the EKF may perform more robustly with sparse measurements at the cost of added testing and user-interfacing with the code. In many settings, these *ad hoc* parameters are tuned by the user to achieve optimal performance.

Contrary to the original EKF, the PC-EKF maintains physical consistency due to its bounded uncertainty terms, leading to a wider allowable measurement frequency range before the measurements are too sparse to obtain reasonable results. The robustness of the PC-EKF is demonstrated, i.e., there are no failed simulations with a wide range of initial conditions and Kalman update frequencies since the physics and mathematical constraints are unconditionally satisfied. This is especially significant for physics applications where low-frequency data acquisition may lead to sparse data for high-frequency phenomena of interest. Since the PC-EKF parameters, P_0 and $P_{\min}$, are not optimized for the system, this discussion does not include a detailed analysis of the accuracy of the estimates. Overall, the time-averaged $\hat{\sigma_L}$ showed good agreement with the true σ_L, completing the verification of the EKF model to a dynamic problem.

3.2 Driven-Damped Harmonic Oscillator

Further verification of the physics-constrained extended Kalman filter (PC-EKF) is performed using a simple nonlinear problem to illustrate its ability compared to analytical theory. Here, the driven-damped harmonic oscillator is employed [101]. This type of global model tracks the steady-state oscillation of one or more states based on a dynamic input to the system. The harmonic oscillator problem is widely used because it represents a system that will experience a restoring force when displaced from its equilibrium position. The driven-damped description indicates that the system studied includes a frictional force which damps the oscillations and would cause the oscillations to decrease over time without the inclusion of a time-dependent driving term to continuously supply energy to the system.

3.2.1 Oscillator Equations

Consider a system that can be written as,

$$\frac{\partial x_1}{\partial t} + \alpha x_1 = x_1 x_2 z, \tag{3.8a}$$

$$\frac{\partial x_2}{\partial t} + \gamma(x_2 - x_{2,int}) = -x_1 x_2 z, \tag{3.8b}$$

where x_1 and x_2 are the states, t is the time, α and γ are constants, and $x_{2,int}$ is the constant value of the x_2 quantity. The driver of the system, z, is assumed to be:

$$z = z_0 \left[1 + A\cos(\omega_r t)\right], \tag{3.9}$$

where z_0 is the mean value of the oscillation, A is the amplitude of the oscillation, and ω_r is a constant driving frequency.

3.2.2 Simulation Setup

This work uses the system described above with the developed physics-constrained extended Kalman filter. The filter is run without *a priori* measurement noise, i.e. $\sigma_R = 0$, to demonstrate a clean, baseline condition to assess the numerical sensitivities of the filter given the available analytical analysis that can be performed with harmonic oscillators. Note that the previous case, the Lorenz system, includes a random measurement noise.

Studies using multistep neural networks to solve supervised learning tasks while respecting physical laws given by nonlinear partial differential equations [102]. Mendible et. al. have used dynamic mode decomposition and a deep Koopman embedding to study rotating detonation waves whose dynamics can, in part, be described using harmonic oscillators [103].

The state vector given to the filter is written as $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & z \end{bmatrix}^T$. To demonstrate the capability of the PC-EKF, $x_{1,\text{ref}}(t)$ is used as the measurement data $\tilde{\boldsymbol{y}}$, from which the estimate of the input parameter $\hat{z}(t)$ is obtained as well as the state estimations of $\hat{x}_1(t)$ and $\hat{x}_2(t)$. This allows for an assessment of the errors associated with the state estimation with respect to the reference solutions, $\epsilon = q_{\text{ref}} - \hat{q}$, where $q = x_1, x_2, z$. The unknown z state is considered piecewise constant for the propagation phase, indicating that it will only change due to the Kalman update.

The reference solution is generated using an initial condition of $(x_1, x_2, z) = (0.009, 0.1, 725)$, constants $\alpha = 100$ and $\gamma = 1$, a frequency of $\omega_r = 2\pi$, and $x_{2,int} = 1$. A small amplitude of $A = 0.02$ is chosen with $z_0 = 725$. The continuous prediction of the EKF is performed using a fourth-order Runge-kutta scheme for a normalized time of twenty units with a time step size of 0.001.

The measurement data are used in the EKF with different sampling frequency to investigate the effects of the data acquisition quality on the EKF estimation. For these results, the initial conditions of the estimated $\hat{x}_1$, $\hat{x}_2$, and $\hat{z}$ quantities are $\hat{x}_1 = 0.0085$, $\hat{x}_2 = 0.08$, and $\hat{z} = 800$. The corresponding initial variance values are set to $(P_{11}, P_{22}, P_{33}) = (1, 1, 4)$ while $P_{ij} = 0$ for $i \neq j$.

3.2.2.1 Physics-Constrained EKF Setup

From Eqs. (3.8a) and (3.8b), the Jacobian matrix for the continuous propagation of the system covariance can be written as

$$F(\boldsymbol{x}(t), t) = \begin{bmatrix} x_2 z - \alpha & x_1 z & x_1 x_2 \\ -x_2 z & -x_1 z - \gamma & -x_1 x_2 \\ 0 & 0 & 0 \end{bmatrix}, \tag{3.10}$$

where the unknown parameter, z, is assumed piecewise constant during the continuous propagation stage, and therefore has no derivatives.

As in Sec. 3.1, the process noise covariance is not going to be set as a constant value. Instead, for this three-equation system, the variance of the unknown state, $P^-_{mm,k+1}$ where $m = 3$ corresponds to the desired unknown driving parameter z is set as a constant, P_0. For the measurement noise covariance, R is designed to be adjusted as follows:

$$R \geq \frac{\left(P^-_{13}\right)^2}{P^+_{33} - P_{\min}} - P^-_{11}, \tag{3.11}$$

where the $P_{\min}$ is a user-defined value for the minimum allowed unknown state covariance which increases R to keep $P^+_{33} > 0$. Note that this minimum variance value is imposed on the variance of the unknown state, P_{mm}, where m is the index corresponding to the unknown state, i.e. $\min(P^+_{mm}) = P_{\min}$. The constant variance value from Eq. (2.18) is set as $P_0 = 95,000$. Although it was not used in this testcase, the minimum P_{mm} value, $P_{\min}$, where $m = 3$, is set to 0.1 for the simulation.

3.2.3 Results

Figure 3.5 shows the reference solution obtained by solving the driven-damped harmonic oscillation equations, Eqs. (3.8) and (3.9), overlaid with the estimation results and 3σ variance bounds obtained from the PC-EKF using a measurement frequency of 250 Hz. The results presented in

this figure correspond to the corrected (+) values of the filter. Due to the lack of added noise to the measurement vector, the measurement noise variance remains zero, $R = 0$, for the entirety of the estimation process. The lack of noise simplifies this test case to investigate the PC-EKF solely depending on the quality of the data acquisition while also further reducing the uncertainty bounds of the states. The transient time for the EKF to achieve a state estimate synchronized with the actual states lasts only a few oscillation cycles. As with the Lorenz system estimate for the unknown σ_L state, the unknown z uncertainty bounds are shown to remain constant throughout the simulation due to the P_0 set by the PC-EKF.

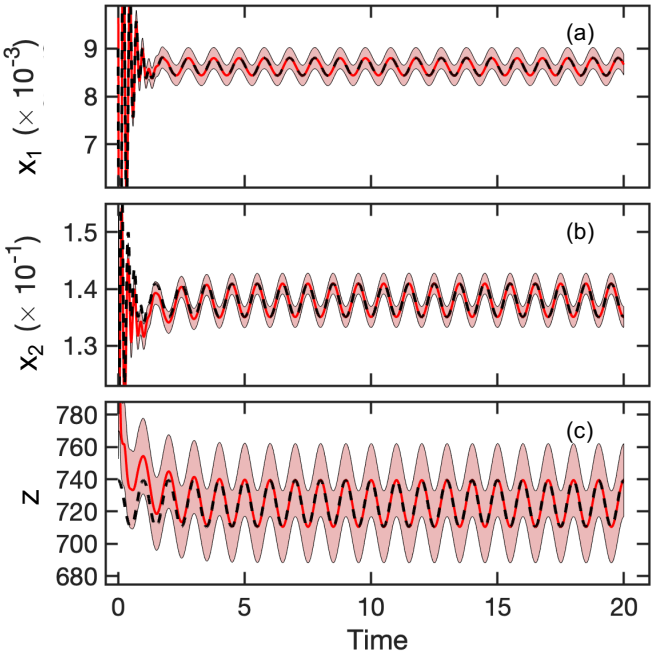

Figure 3.5: An example of the reference solution, $\boldsymbol{x}_{\text{ref}}$, in black dashed line, overlaid with the updated, (+), estimates of the physics-constrained extended Kalman filter (PC-EKF) estimation solution, $\hat{\boldsymbol{x}}$, in the red solid line. The 3σ variance bounds are shown in the pink shade. These figures present (a) x_1, (b) x_2, and (c) z.

Since the reference solutions for all quantities in the system are pre-calculated and unaffected by noise, Fig. 3.6 shows the system error (e.g. residual) at every point in time calculated as the difference between the reference solution and the most recent estimation. A 3σ bound, where σ

is equivalent to the standard deviation, $\sigma_i = (P_{ii}^+)^{1/2}$ for $i = 1, 2, 3$, is used both to generate a region of confidence for the estimation and to set a variance bound on the error. It can be seen from Fig. 3.6 that the error remains bounded for all time, even during the initial transient time of the system, for both of the unknown estimates, $\hat{x}_2$ and $\hat{z}$. Note that because the measurement noise is set to zero, the variance of the $\hat{x}_1$ estimate, i.e. P_{11}, becomes zero after the first Kalman update. This is because the Kalman gain for the x_1 state, i.e. K_1 in Eq. (3.4), becomes 1, indicating that the EKF model fully accepts the measurement as $\hat{x}_1$. Despite the initially large set variance in $\hat{z}$, the standard deviation settles to 23, the standard deviation of the $\hat{x}_2$ estimate quickly settles to 0.0022 which can be seen as a very narrow bound.

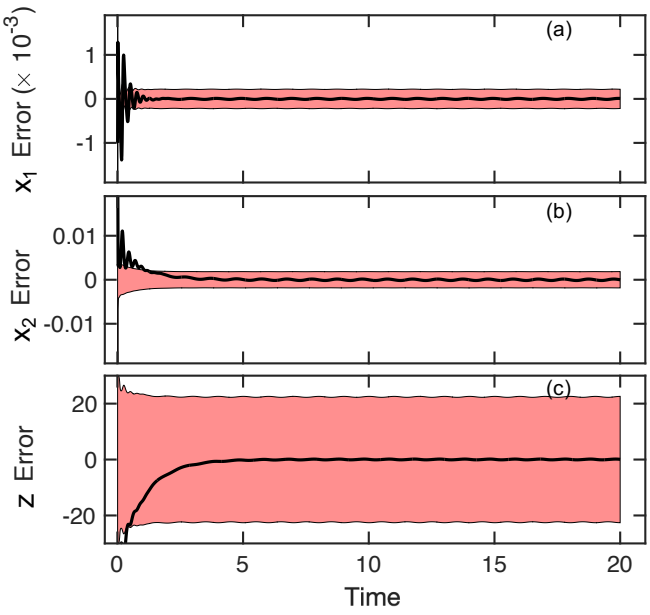

Figure 3.6: The corresponding residual error (black solid) for each of the three updated state values, i.e. $\hat{\boldsymbol{x}}^+$, plotted with their 3σ bounds (pink shade) for (a) x_1, (b) x_2, and (c) z.

3.2.3.1 Comparison against Theory

Having determined that the PC-EKF can estimate the driven-damped harnonic oscillator within narrow error bounds given sufficient measurement data and reasonable initial conditions, an ana-

lytical analysis of the filter results is now presented. Because a small A value is chosen, the harmonic oscillation remains sinusoidal throughout the simulation. Under this condition, the linear perturbation theory can be used to obtain the exact solution for the dynamical system, allowing for a *verification* of the EKF model. Characteristics specific to this system include the amplitude and phase of the x_1-, x_2-, and z-state oscillations as calculated from driven-damped harmonic oscillator theory and the estimation results [104]. For a larger A value, the oscillation becomes non-sinusoidal, i.e. the dynamics are nonlinear, making the linear perturbation analysis invalid.

Using z_0, the equilibrium (time-averaged) values of the x_1 and x_2 states (denoted by subscript 0) can be calculated as

$$x_{1,0} = -\frac{\gamma}{\alpha}\left(\frac{\alpha}{z_0} - x_{2,int}\right), \tag{3.12a}$$

$$x_{2,0} = \frac{\alpha}{z_0}. \tag{3.12b}$$

Applying a linear perturbation theory of the form $v = v_0 + v'$ to both state variables, i.e. v can be either x_1 or x_2, in conjunction with the equilibrium solution from Eqs. (3.12a) and (3.12b) leads to two second-order equations of similar form. The general second-order equations can be written as

$$\ddot{v} + \nu_d \dot{v} + \omega_0^2 v = \omega_0^2 \left[M_{0,v}\cos(\omega_r t) + N_{0,v}\sin(\omega_r t)\right], \tag{3.13}$$

where $\nu_d = \gamma z_0 x_{2,int}/\alpha$ is the damping frequency, $\omega_0 = \left(\gamma z_0 x_{2,int} - \gamma\alpha\right)^{1/2}$ is the natural (harmonic) frequency of the system, and $M_{0,v}$ and $N_{0,v}$ are the amplitude of the forced oscillation, which have a different form depending on the state variables. For the x_1-state, these forced oscillation amplitudes are $M_{0,x_1} = A\gamma/z_0$ and $N_{0,x_1} = -A\omega_r/z_0$. For the x_2-state, they are $M_{0,x_2} = -A\alpha/z_0$ and $N_{0,x_2} = A\omega_r/z_0$.

Given the second-order differential equation, assuming a linear perturbation, presented in Eq. (3.13), the amplitude and phase can be written as

$$v_{\mathrm{AMP}} = \frac{\omega_0^2}{\sqrt{(\omega_0^2 - \omega_r^2)^2 + (\nu_d \omega_r)^2}} \sqrt{M_{0,v}^2 + N_{0,v}^2}, \tag{3.14}$$

and

$$\tan \phi_v = \frac{M_{0,v}\nu_d\omega_r + N_{0,v}(\omega_0^2 - \omega_r^2)}{M_{0,v}(\omega_0^2 - \omega_r^2) - N_{0,v}\nu_d\omega_r}, \tag{3.15}$$

respectively. Here, v_{AMP} is the amplitude and ϕ_v is the phase shift angle between the select quantity $v = x_1, x_2$ and the driving function z. The amplitude and phase of the estimated variables are calculated by taking the Fourier transform. To study the effect of different data acquisition frequencies, some user-defined number of data points are skipped during the creation of the different measurement datasets. The measurement intervals are kept equidistant, i.e. constant $\Delta t = t_{k+1} - t_k$, for simplicity. Increasing intervals between measurements corresponds to decreasing the acquisition frequency where Kalman updates are performed since the measurement frequency is $1/\Delta t$. The highest measurement frequency corresponds to the case where measurement updates are performed at every time step used to generate the reference solution. For this study, the reference solution is obtained using a time step that resolves one oscillation cycle with 1000 time steps.

Table 3.1 shows the analytical results of the oscillation amplitude for various data acquisition frequencies. Since the Kalman gain has $K_k = 1$ due to $R_k = 0$, the amplitude of the estimate $\hat{x}_1^+$ should be identical to the reference solution $x_{1,ref}$. The error of the amplitude for x_1-state originates from the error associated with the amplitude calculation using a fast Fourier transform (FFT). It can be seen that the x_1 amplitude error is less than 0.1 % regardless of the sparsity of the measurement data.

Given the errors associate with the FFT calculation, it is remarkable to see that the error of the $\hat{x}_2$ and $\hat{z}$ estimate amplitudes are relatively small ($\varepsilon < 5$ %) for the range of measurement frequencies studied. Once the frequency drops to 10 samples per oscillation or fewer, the PC-EKF does not have enough resolution in the measurement to fully capture the system dynamics and the error begins to increase as the frequency decreases. It can therefore be recommended that 20-40

measurement data per cycle are needed to capture this oscillation signal with high accuracy. A negative sign on the relative error in Table 3.1 indicates that the estimate has a larger amplitude than the reference value.

Table 3.1: Sensitivity study of sparse data by comparison of the reference and estimated amplitude of the oscillation with their relative errors. The reference solutions are given in the parenthesis of the column headers. The relative errors are calculated as $(x_{1,exact} - \hat{x}_1)/x_{1,exact}$.

Measurement Frequency	$\boldsymbol{x_{1,AMP}}$ **(1.86780×10^{-4})**	(ε_{x_1})	$\boldsymbol{x_{2,AMP}}$ **(2.9415×10^{-3})**	(ε_{x_2})	$\boldsymbol{z_{AMP}}$ **(14.5)**	(ε_z)
1000	1.8680×10^{-4}	(-0.011%)	2.9418×10^{-3}	(-0.009%)	14.4961	(0.027%)
500	1.8680×10^{-4}	(-0.011%)	2.9418×10^{-3}	(-0.009%)	14.4961	(0.027%)
250	1.8680×10^{-4}	(-0.011%)	2.9413×10^{-3}	(0.008%)	14.4964	(0.025%)
200	1.8680×10^{-4}	(-0.011%)	2.9413×10^{-3}	(0.008%)	14.4966	(0.023%)
125	1.8680×10^{-4}	(-0.011%)	2.9418×10^{-3}	(-0.009%)	14.4975	(0.017%)
100	1.8680×10^{-4}	(-0.011%)	2.9418×10^{-3}	(-0.009%)	14.4984	(0.011%)
80	1.8671×10^{-4}	(0.038%)	2.9426×10^{-3}	(-0.037%)	14.5036	(-0.025%)
50	1.8680×10^{-4}	(-0.011%)	2.9417×10^{-3}	(-0.006%)	14.5051	(-0.035%)
40	1.8680×10^{-4}	(-0.011%)	2.9414×10^{-3}	(0.004%)	14.5097	(-0.067%)
20	1.8679×10^{-4}	(0.005%)	2.9364×10^{-3}	(0.174%)	14.5326	(-0.225%)
10	1.8679×10^{-4}	(0.005%)	2.8210×10^{-3}	(4.098%)	14.2168	(1.953%)

The $\hat{x}_2$ and $\hat{z}$ results can provide better insight into the capabilities of the PC-EKF. The results shown in Figs. 3.7(a) and 3.7(b) are a proof of the convergence of the EKF model to the exact solution, in the absence of measurement noise. This illustrates that such state estimation methods can be of use to understand physical phenomena. The errors of $\hat{x}_2$ and $\hat{z}$ do not show a monotonic trend, as shown in Figs. 3.7(a) and 3.7(b), but the absolute values of the errors are within 0.3% of the reference solution when 20 or more measurement updates are performed per oscillation period. The overestimation seen at some points in both of the unobserved states is likely due to the piecewise constant assumption for the $\hat{z}$ estimate in the present PC-EKF model. The z_{ref} is a continuous time-dependent function, while $\hat{z}$ is considered to be a state (parameter) that is constant in time between the measurement updates in the EKF model. The constant $\hat{z}$ during the continuous update results in a phase lag for the estimate $\hat{x}_1$ and $\hat{x}_2$ in comparison to $z(t)$ and leads

to a discrepancy in the amplitude estimation.

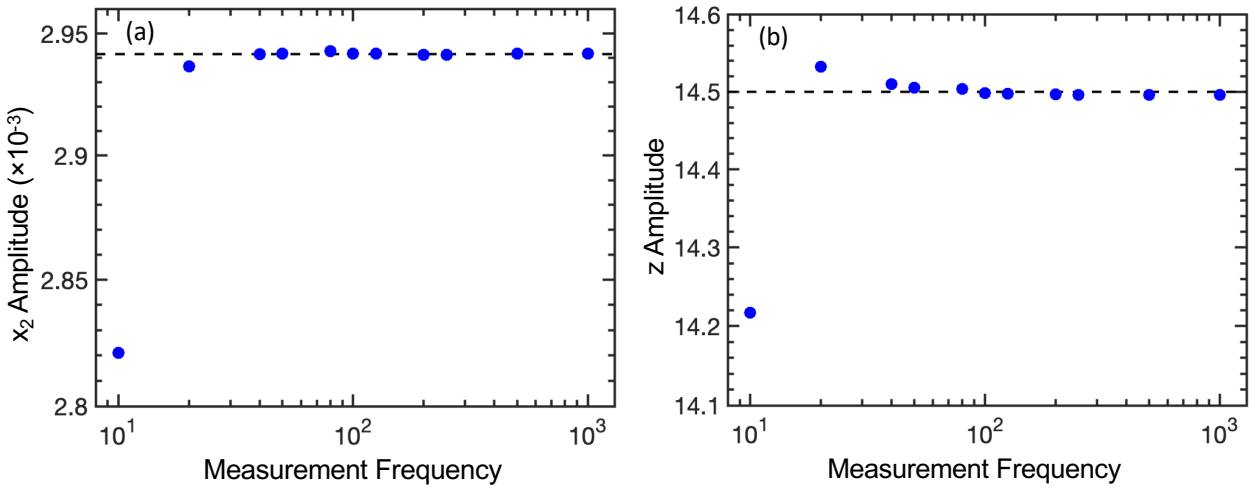

Figure 3.7: (a) The estimated $\hat{x}_2$-amplitude (circle) of the driven-damped harmonic oscillator as it is affected by the data acquisition, i.e. measurement, frequency of the PC-EKF compared to the theoretical solution (dashed line). Convergence is observed such that as more detailed measurement signals are given to the filter, the original physics are more clearly resolved. (b) The estimated $\hat{z}$-amplitude (circle) of the driven-damped harmonic oscillator compared to the theoretical solution (dashed line).

Table 3.2 shows the phase of the oscillations for the same measurement frequencies used to study the amplitude. Since the phase is calculated from 0 to 2π, the absolute error of the phase lag, $\varepsilon = q_{\text{exact}} - \hat{q}$ where $q = x_1, x_2, z$, is used instead of the relative error. The phase lag $\phi_{21} = \phi_2 - \phi_1$ denotes the difference between the $\hat{x}_2$ and $\hat{x}_1$ phases where ϕ_2 and ϕ_1 are the phase lag of $\hat{x}_2$ and $\hat{x}_1$ with respect to the input function, $\hat{z}$. Similar to the $\hat{x}_1$ amplitude results, the phase of the $\hat{x}_1$ signal is almost identical to that of the reference solution. The $\hat{x}_2$ and $\hat{z}$ quantities exhibit near-monotonically changing values dependent on the measurement frequency. As both the $\hat{x}_1$ and $\hat{x}_2$ phases are calculated with respect to the $\hat{z}$ phase, the phase lag between the $\hat{x}_2$ and $\hat{x_1}$ phases is shown to more clearly demonstrate the data trends.

Figure 3.8 shows the absolute error of the phase lag between $\hat{x}_2$ and $\hat{x}_1$ and the $\hat{z}$-state phase from Table 3.2. When the frequency of data acquisition, i.e. Kalman update, increases, the absolute

Table 3.2: Sensitivity study of sparse data by comparison of the reference and estimated phase angles given in radians. The reference solutions are provided in parenthesis in the column headers. The input z_{ref} is a cosine function shown in Eq. (7), thus the exact phase lag is 0. The exact phase lags for x_1 and x_2 are derived from Eq. (3.15).

Measurement Frequency	$\boldsymbol{\phi_1}$ **(-1.3353)**	$\boldsymbol{\phi_2}$ **(0.01490)**	$\boldsymbol{\phi_z}$ **(0)**	$\boldsymbol{\phi_{21} = \phi_2 - \phi_1}$ **(1.350215)**	$(\varepsilon_{\phi_{21}})$
1000	-1.3353	0.01494	1.84×10^{-3}	1.350218	(-3.00×10^{-6})
500	-1.3353	0.01494	3.64×10^{-3}	1.350217	(-2.00×10^{-6})
250	-1.3353	0.01494	7.22×10^{-3}	1.350212	(3.00×10^{-6})
200	-1.3353	0.01493	9.00×10^{-3}	1.350206	(9.00×10^{-6})
125	-1.3353	0.01490	1.43×10^{-2}	1.350172	(4.30×10^{-5})
100	-1.3353	0.01486	1.78×10^{-2}	1.350130	(8.50×10^{-5})
80	-1.3350	0.01478	2.12×10^{-2}	1.349826	(3.89×10^{-4})
50	-1.3353	0.01426	3.44×10^{-2}	1.349533	(6.80×10^{-4})
40	-1.3353	0.01362	4.22×10^{-2}	1.348892	(1.39×10^{-3})
20	-1.3353	0.00453	7.29×10^{-2}	1.339796	(1.04×10^{-2})
10	-1.3353	-0.53921	5.52×10^{-2}	1.281367	(6.88×10^{-2})

error of the phase lag between $\hat{x}_2$ and $\hat{x}_1$ is shown to approach zero. This indicates that the $\hat{x}_1$ and $\hat{x}_2$ phases are individually approaching their exact phases with regards to the reference solution (see Table 3.2). In particular, the phase of $\hat{z}$ and the phase lag between the $\hat{x}_1$ and $\hat{x}_2$ states, ϕ_{21}, show a nearly monotonic trend, approaching the exact solution which can be obtained from Eq. (3.13). By taking the log-log plot, the slope of the errors denote the order of accuracy, i.e. $\varepsilon \propto (\Delta t)^p$. The slopes of ϕ_{21} and ϕ_z are estimated to be $p = 3.05 \pm 0.13$ and $p = 0.93 \pm 0.017$, respectively. All of the phase and amplitude results demonstrate convergence of the PC-EKF as the measurement frequency increases.

As a larger time interval between consecutive data acquisitions is used, the measurement frequency decreases and the correction data becomes more sparse. Furthermore, as the number of data points used to resolve an oscillation decreases, the finer details of the oscillation cannot be captured. This causes the amplitude to decrease and the phase angle to shift. It is observed from the present PC-EKF model that the amplitude of the estimated state becomes smaller as the data acquisition becomes more sparse, although the time-averaged value of the oscillations are still estimated

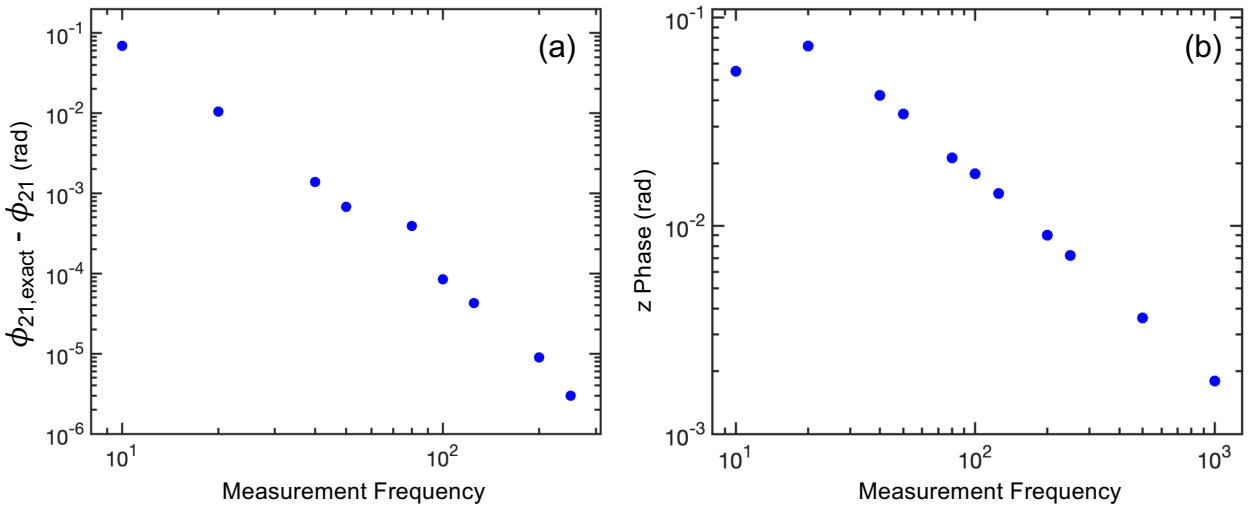

Figure 3.8: (a) The absolute error of the phase lag between $\hat{x}_2$ and $\hat{x}_1$ and (b) the $\hat{z}$ estimate phase for different measurement frequencies. Note that the last two points, i.e. measurement frequency of 500 and 1000 samples per unit time, in Fig. 3.8(a) are not shown since the absolute errors are negative as shown in Table 3.2.

well for all three states. While the high-frequency oscillations may be underresolved, the ability to capture time-averaged values is of significant importance for robust estimation with sparse data. In some cases, physics occurs at small scales that experimental measurement may not be able to resolve it, e.g. electron turbulence, or the particular dynamics of a system are simply too chaotic to be resolved. In these cases, being able to robustly capture even just the time-averaged quantities could be of immense benefit to the scientific community to inform computational models.

3.3 Xenon Oscillations

Having demonstrated the added robustness of the physics-constrained EKF as well as the convergence of the estimates towards the reference solution given sufficient measurement data, the PC-EKF is applied to the time-dependent discharge plasma of a Hall effect thruster. For this study, the discharge current measurement data is used to estimate the electron temperature using two different, simple global model representation of the thruster. The focus of this investigation is to directly demonstrate the capabilities of the PC-EKF for plasma physics research applications.

As mentioned in the introduction, Hall effect thrusters (HETs) are plasma-based electric propul-

sion (EP) devices that can be used for orbit raising and station keeping of in-space satellites. The low-temperature magnetized plasmas in HETs exhibit various types of oscillations due to the multiscale and multiphysics nature of the flow. The plasma flow is inherently three-dimensional and is affected by the complex physical phenomena including plasma instabilities, turbulence, plasma-material interactions, collisions and radiation, and reactions [54]. These include so-called self-organized propagating structures at low-frequencies [105]. While there have been experimental and computational efforts to study these self-organizing structures [106, 107, 108, 109, 41, 30], the mechanisms of such remain poorly understood. Another key physical process is the anomalous electron transport across the magnetic field lines, which determines the resistivity of the plasma flow, affecting the ionization and acceleration mechanisms [8].

Data-driven modeling has begun to permeate the Hall effect thruster community. Recent work by Jorns et. al. has studied the creation of low-dimensional representations of physical models to study plasma phenomena as functional forms [65] while others have looked to forms of mode decomposition [66, 110]. Work performed before this dissertation studied time-embedded phase space plots to optimize input conditions for a Hall effect thruster model [68]. Work has also been performed to study the application of deep recurrent neural networks to supplement PIC simulations to better predict thruster dynamics in a one-dimensional model [111].

3.3.1 0D Ionization Oscillation Model - 3 Equation

To simplify the plasma oscillations that are driven by the electron dynamics, a zero-dimensional (0D) global model can be considered that accounts for the domain as a single spatial unit. The time-dependent volume-averaged number densities can be solved, taking into account various volumetric reactions (e.g., ionization, recombination, etc.) and particle fluxes to the surface as loss mechanisms. For this work, singly charged ions and direct ionization from ground state to the ionized state are considered. The interaction between ion and neutral number densities is captured as a function of time driven by the time-dependent electron temperature [112], using a set of first-order nonlinear differential equations that describe the interaction between two entities, cf. Lotka-Volterra model [113, 114]. In Ref. 112, it was proposed that the plasma oscillations are

unconditionally damped in the absence of any electron temperature fluctuation. While the ionization oscillations may be due to other mechanisms, such as the 1D advective nature of the plasma flow [115, 116], the 0D global model shows that one possible mechanism for the ionization oscillation is for the electron temperature to vary in time, similar to a driven-damped harmonic oscillator. The true physical system is affected by multidimensional processes which could also be coupled with an EKF to investigate plasma phenomena such as electron transport [28].

The volume-averaged ion and neutral continuity equations can be written as

$$\frac{dN_i}{dt} + \frac{N_i U_i}{L_{ch}} + \frac{2N_i U_{i,w}}{R_\triangle} = N_i N_n \xi_{\text{ion}}, \tag{3.16a}$$

$$\frac{dN_n}{dt} + \frac{(N_n - N_{int})U_n}{L_{ch}} = -N_i N_n \xi_{\text{ion}}, \tag{3.16b}$$

where *N* is the number density, *U* is the mean velocity, subscripts *i* and *n* denote ion and neutral atoms, respectively, N_{int} is the number density of neutral particles injected from the anode, $U_{i,w} = (eT_e/M_i)^{1/2}$ is the ion acoustic speed which contributes to ion diffusion to the wall, T_e is the electron temperature in electron-volts, M_i is the mass of an ion, L_{ch} is the channel length, and $R_\triangle$ is the channel width. Here, $U_i = (eV_D/M_i)^{1/2}$, where V_d is the discharge voltage and e is the elementary charge. The ionization rate coefficient ξ_{ion}, whose unit is m^3/s, is given as

$$\xi_{\text{ion}} = \left[AT_e^2 + B\exp\left(-\frac{C}{T_e}\right)\right]\left(\frac{8eT_e}{\pi m_e}\right)^{1/2}, \tag{3.17}$$

where $A = -1 \times 10^{-24}$, $B = 6.386 \times 10^{-20}$, $C = 12.13$, and m_e is the electron mass [4].

3.3.2 Simulation Setup

This testcase uses the measurement data of the discharge current oscillation of a SPT-100 Hall effect thruster. The time-dependent discharge current of several operating modes of the SPT-100 is measured by Gascon, Dudeck, and Barral in Ref. 3. The operation metrics for each mode are given in detail in the referenced work, as well as the defining characteristics of each operation

mode, which vary with mass flow rate, discharge voltage, and magnetic field strength [117, 118]. For this test case, four different discharge oscillation modes are chosen from Ref. 3 to estimate the electron temperature using the PC-EKF model. Their relevant operating conditions are provided in Table 3.3.

Table 3.3: The initial conditions used for the filter estimate for each of the four test cases taken from Ref. 3. Cases I, II, and IV use borosil for the channel walls while case III uses alumina.

Case	**I**	**II**	**III**	**IV**
$V_d(V)$	70	125	200	275
$\langle I_d \rangle$	2.73	4.42	4.82	3.97
$\hat{N}_i(\text{m}^{-3})$	1.5×10^{17}	1.5×10^{17}	1.5×10^{17}	1.5×10^{17}
$\hat{N}_n(\text{m}^{-3})$	9×10^{18}	9×10^{18}	7×10^{18}	7×10^{18}
$\hat{T}_e(eV)$	18	18	30	30

For the EKF, the state variables are taken as $\boldsymbol{x} = \begin{bmatrix} N_i & N_n & T_e \end{bmatrix}^T$. Note that T_e is considered to be constant during the continuous prediction stage, but is corrected at every measurement update. Here, time-dependent ion number density data, N_i, are calculated from the discharge current and used as the measurement data, $\tilde{\boldsymbol{y}}$. The ion number density is calculated by assuming a beam utilization efficiency, which is the ratio of ion current, I_i, to the total discharge current I_d, as $\eta_b = I_i/I_d$. The measurement data can be obtained as $\tilde{\boldsymbol{y}} = N_i = \eta_b \tilde{I}_d/(U_i e A_c)$, where A_c is the thruster channel area, which can be calculated from the thruster geometry using $L_{ch} = 2.5$ cm and $R_\Delta = 1.5$ cm, and $\tilde{I}_d$ is the measured discharge current. As discussed in Ref. 3, the discharge current is measured at a rate of 200 kHz. It is to be noted that the 0D model assumes a unitary value for the plasma properties, while the plasma is inhomogeneous in the actual HETs. Thus, while $\eta_b = 0.12 - 0.2$ is used for this present study, these values do not match other literature. This is considered acceptable for this dissertation because the purpose of this study is to demonstrate the capability of the EKF model rather than obtain perfectly accurate estimates.

The initialized value for the estimated ion number density, $\hat{N}_i$, is taken from the mean value of the measurement data. We can make educated guesses for the other states based on this ion

number density and assumed operating conditions of the thruster. The initial variances, P_{mm}, are set assuming that the initial density estimate is inaccurate. Here, $P_{mm} = (1 \times 10^{34}, 1 \times 10^{36}, 10)$ is considered for the initial condition. A data acquisition frequency of 200 kHz is assumed as that is the rate at which the original data was obtained, although some step sizes are not constant within each solution. The mass flow rate is 5 mg/s and the average neutral velocity is assumed to be 400 m/s based on previous studies in Ref. 119 that show neutral atom velocity acceleration by the channel exit.

3.3.2.1 Physics-Constrained EKF Setup

Based on the state vector described above, the Jacobian matrix for this system of equations is written as,

$$F(\boldsymbol{x}(t), t) = \begin{bmatrix} -U_i/L_{ch} - 2U_{i,w}/R_\Delta + N_n\zeta_{ion} & N_i\zeta_{ion} & -2/R_\Delta \partial U_{i,w}/\partial T_e + N_n N_i \partial\zeta/\partial T_e \\ -N_n\zeta & -U_n/L_{ch} - N_i\zeta & -N_n N_i \partial\zeta/\partial T_e \\ 0 & 0 & 0 \end{bmatrix}, \tag{3.18}$$

where the partial derivatives of $U_{i,w}$ and ζ with respect to T_e can be written as

$$\frac{\partial U_{i,w}}{\partial T_e} = \frac{1}{2}\sqrt{\frac{e_{ch}}{T_e M_i}}$$

and

$$\frac{\partial \zeta}{\partial T_e} = \sqrt{\frac{8e_{ch}T_e}{\pi m_e}}\left(2AT_e + BC\frac{\exp(\frac{-C}{T_e})}{T_e^2}\right) + \frac{1}{2}(AT_e^2 + B\exp(\frac{-C}{T_e}))\sqrt{\frac{8e_{ch}}{T_e \pi m_e}}$$

.

The noise covariance matrices are initialized as $Q = 1 \times 10^7$ and $R_{k=0} = 1 \times 10^{15}$. For the PC-EKF, a minimum covariance value is set to $P_{\min} = 0.1$ and a fixed variance for the input variable is fixed at $P_0 = 18$ for P_{33}^-, the variance of the unknown state variable, (see Eq. (2.18)) to ensure

that the state estimates satisfy the physical and mathematical constraints.

3.3.3 Results

Figures 3.9 and 3.10 depict the measurement signal and corresponding estimated state time histories for four different thruster discharge modes, $V_d = 70$, 125, 200, and 275 V, to demonstrate the robustness of the PC-EKF at handling different frequencies and amplitudes of plasma oscillations. All cases run without any failure using the PC-EKF for a variety of initial conditions and data acquisitions. Despite faster data acquisition rates, the estimation results were unchanged, and so are not presented within this dissertation. The estimates of the neutral number density and electron temperature exhibit trends that are consistent with the physics. For instance, a decrease in neutral atom density follows an increase in ion density when an ionization event occurs. This ionization is driven by the increase in electron temperature, which can be seen in all four cases. Previous work in Ref. 104 discusses the phase lag between the neutral, plasma, and electron temperature signals. Note that when using the original EKF with fixed *ad hoc* noise covariances, some cases fail without meticulous attention to detail in initializing the filter, particularly for the cases where the oscillation amplitude is large and the time between measurement acquisitions vary. As the robustness of the PC-EKF in comparison to the original EKF is discussed in Sec. 3.1, here, only the results with PC-EKF are presented.

Case A, as shown in Fig. 3.9(a), is the operation mode at very low discharge voltage, $V_d = 70$ V. Gascon et al. showed that the operation mode below this voltage is often unsteady [3]. In fact, the neutral number density, shown in Fig. 3.9(a3), exhibits a relatively repeatable oscillation up to 0.6 ms, where the neutral number density then increases despite the same level of ion density at $t > 0.6$ ms. This transition is due to the incoherent ion density fluctuation, shown in Fig. 3.9(a2) after 0.6 ms where the signal becomes more akin to a steady-state value perturbed by noise than being dominated by a low frequency, low amplitude oscillation. The PC-EKF model provides intriguing insights in terms of the change in electron temperature required to sustain the plasma. At $t < 0.6$ ms, the time-averaged T_e is approximately 18 eV, while it becomes $16 - 17$ eV at $t > 0.6$ ms as the discharge current trace amplitude reduces. The required time-averaged electron

temperature is higher in the presence of a low-frequency oscillation and gives evidence that the PC-EKF is able to detect and estimate mode transitions based on a discharge current oscillation trace. This is a significant ability as mode transitions have been studied in detail with computational models [8]. The ability to use data-driven modeling to gain additional insight into mode transitions would be invaluable to the community, where much data-driven modeling to date has focused on steady-state operation conditions.

Figure 3.9(b) shows Case B where larger-amplitude oscillations are observed at very low frequencies (5 kHz) that can be considered to 'loop' or 'circuit' oscillations [29], caused by effects of the thruster circuitry. Due to the large amplitude of the oscillations, the original EKF with fixed measurement and process noise covariances struggled to return good estimates of the system. Either the system estimates oscillate wildly outside of the variance bounds or the filter estimate diverges before the simulation was complete. It is possible that there exists some perfect combination of initial variance values and initial conditions that enable the original EKF to estimate this system well, but part of this research is to develop a robust model that does not require such fine turning. In comparison, the PC-EKF captures state estimates that are consistent with the physics. The electron temperature (see Fig. 3.9(b4)) is highly nonsinusoidal and consequently the neutral atom density (see Fig. 3.9(b3)) decreases substantially by the burst of the ionization. The PC-EKF is able to estimate a nonlinear oscillatory mode where the oscillations are multi-frequency and non-sinusoidal, which is a promising feature for the plasma oscillation studies.

Figure 3.10(a) shows Case C where the breathing mode oscillations in the range of 20-25 kHz are present. The nonsinusoidal oscillation comes from the moderately large-amplitude oscillation of the electron temperature, which is considered the forcing, or driving, function of this system. Coherent oscillations in ion density, neutral atom density, and electron temperature are observed in this operation mode. These large fluctuations cause issues for the original EKF, similar to those in Case B, where the filter returns poor estimates of a few oscillation peaks, exceeding beyond the variance bounds, before ultimately failing. Conversely, the PC-EKF is shown to remain within its confidence bounds and nearly capture the dynamics of the discharge current. The fluctuations of

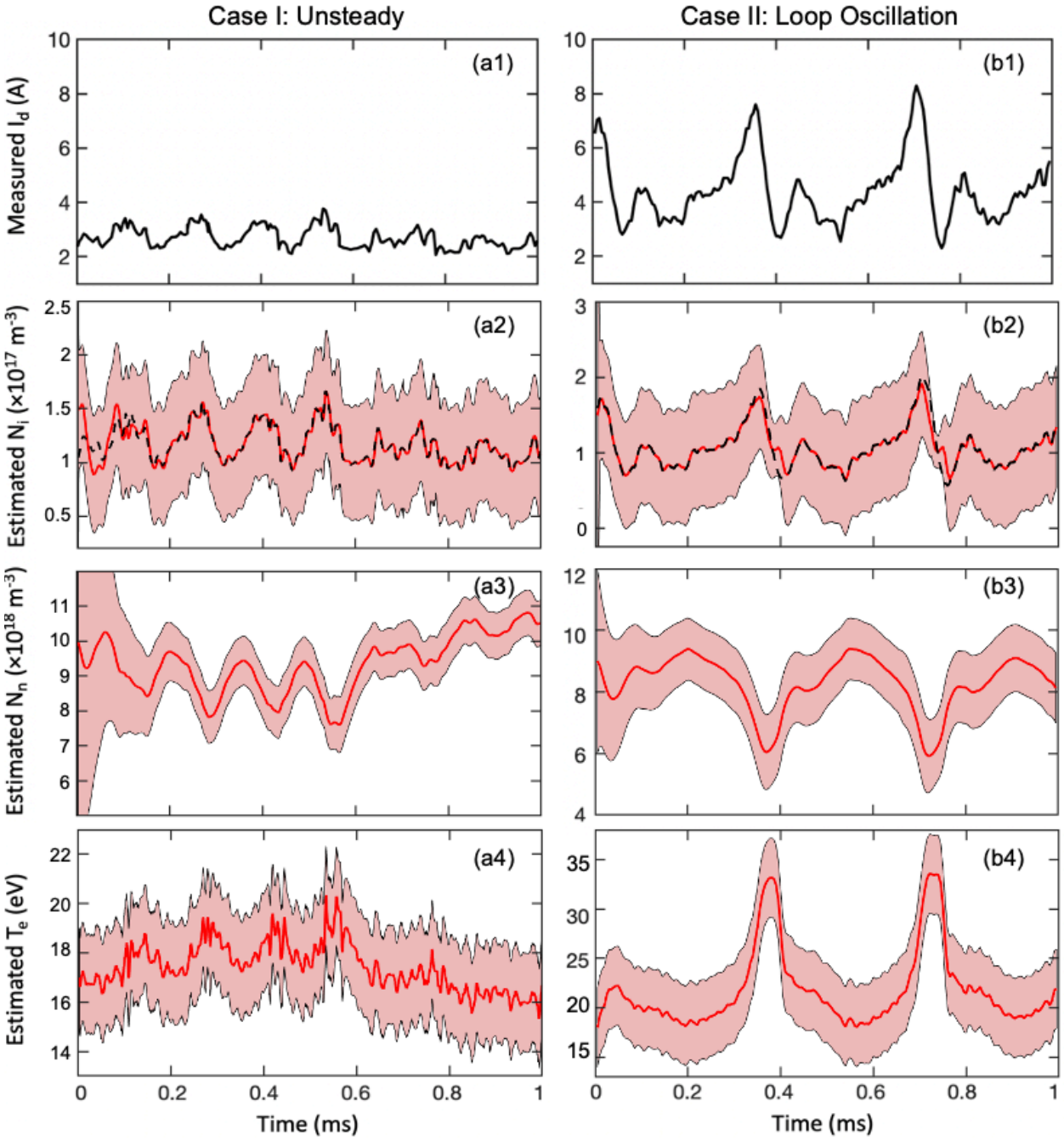

Figure 3.9: Results of the constrained EKF using the 0D ionization oscillation model for various Hall effect thruster operation modes. Here, (a) $V_d = 70$ V, Case I, and (b) $V_d = 125$ V, Case II. The measurement discharge current (a1, b1), estimated ion number density (a2, b2), estimated neutral number density (a3, b3), and estimated electron temperature (a4, b4) are shown, respectively. All estimated values (red) are given with their $\pm 3\sigma$ confidence bounds (pink) except for the electron temperature which is shown with a σ confidence bound. The calculated ion number density measurement is denoted by the black dashed line in the ion number density plots.

the electron temperature are out of phase with the measurement data, lagging by approximately a half phase. Some of this lag can be attributed to the simplicity of the model used to describe these dynamics.

Figure 3.10(b) shows Case D, a near-steady-state operation of the thruster that is mainly marked by random noise. This case does not exhibit a large-amplitude low-frequency oscillation but is not perfectly steady-state either. Such features can be clearly observed in the ion density, neutral density, and electron temperature estimates. Despite the nonsinusoidal higher frequency oscillations in this discharge current trace, the estimates still capture the dynamics well.

3.3.4 0D Ionization Oscillation Model - 6 Equation

Having performed a series of studies using the three equation, predator-prey model formulation of the xenon Hall thruster model, the same measurement signals were estimated using a more complex six equation global model. This work is based on an extended model using Ref. 120 and Ref. 8 that includes three excited states of xenon. A spatially-averaged model that tracks the oscillating number densities of various particle species can be written as,

$$\frac{\partial N_i}{\partial t} - \frac{N_i U_i}{L_{ch}} = k_D(T_e)N_e N_{n,g} + k_{s,m}(T_e)N_{n,m}N_e + k_{s,r}(T_e)N_{n,r}N_e + k_{s,**}(T_e)N_{n,**}N_e, \quad (3.19a)$$

$$\frac{\partial N_{n,g}}{\partial t} + \frac{(N_{n,g} - N_{int})U_n}{L_{ch}} = -k_D(T_e)N_{n,g}N_e - k_{e,m}(T_e)N_{n,g}N_e - k_{e,r}(T_e)N_{n,g}N_e - k_{e,**}(T_e)N_{n,g}N_e + N_{n,r}\nu_{m0}, \quad (3.19b)$$

$$\frac{\partial N_{n,m}}{\partial t} + \frac{N_{n,m}U_n}{L_{ch}} = k_{e,m}(T_e)N_{n,g}N_e + N_{n,**}\nu_{ml} - k_{s,m}N_{n,m}N_e, \quad (3.19c)$$

$$\frac{\partial N_{n,r}}{\partial t} = k_{e,r}(T_e)N_{n,g}N_e + N_{n,**}\nu_{ml} - k_{s,r}(T_e)N_{n,r}N_e - N_{n,r}\nu_{m0}, \quad (3.19d)$$

$$\frac{\partial N_{n,**}}{\partial t} = k_{e,**}(T_e)N_{n,g}N_e - k_{s,**}N_{n,**}N_e - 2N_{n,**}\nu_{ml}, \quad (3.19e)$$

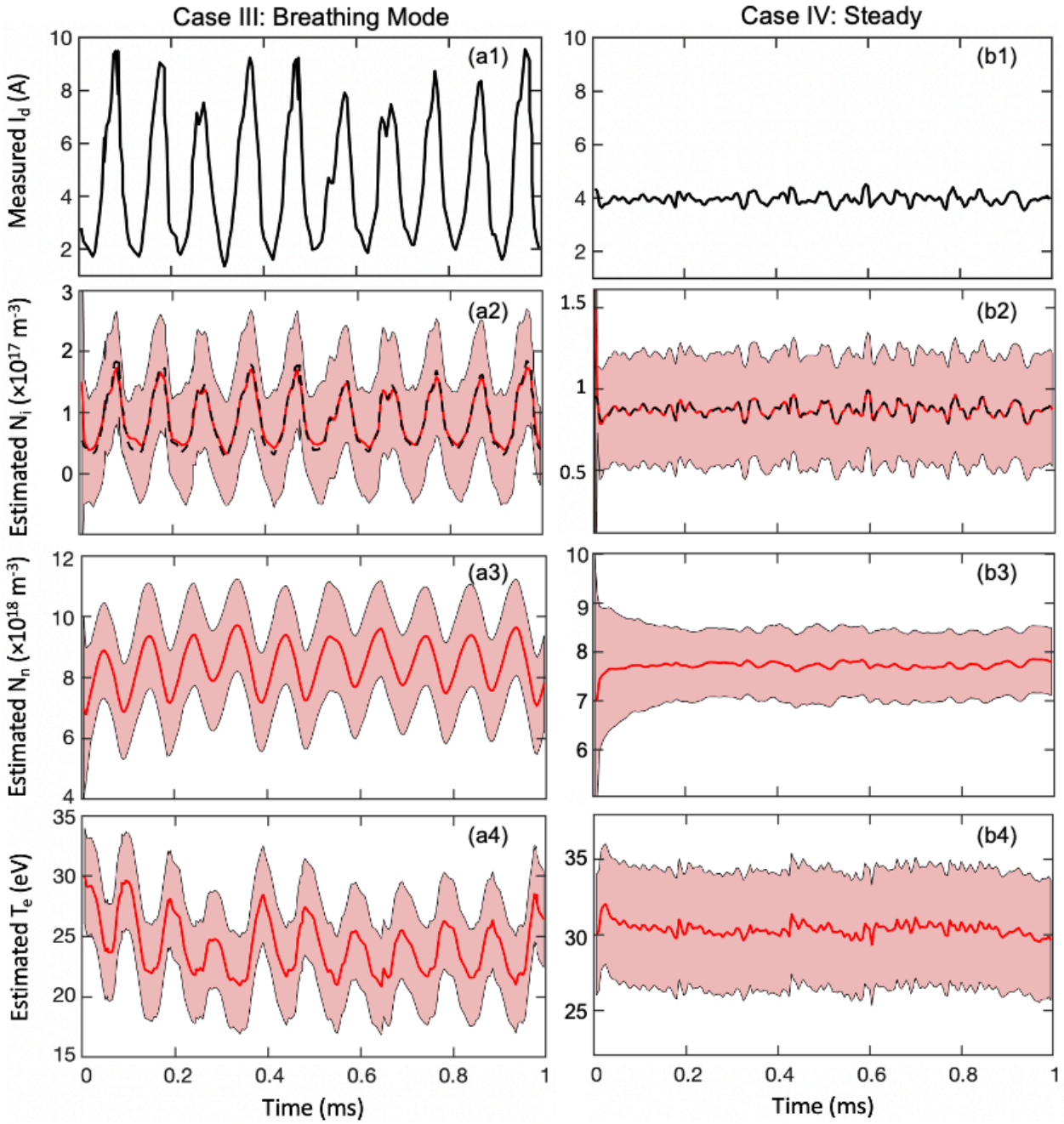

Figure 3.10: Results of the constrained EKF using the 0D ionization oscillation model for various Hall effect thruster operation modes similar to Fig. 3.9. Here (a) $V_d = 200$ V, Case III, and (b) $V_d = 275$ V, Case IV. The measurement discharge current (a1, b1), estimated ion number density (a2, b2), estimated neutral number density (a3, b3), and estimated electron temperature (a4, b4) are shown.

where *N* and *U* are the spatially-averaged number density and mean velocity, the subscripts *i*, *g*, *m*, *r*, and ** denote ion, ground, metastable ($^3P_2$), radiative ($^3P_1$), and higher electronically excited atoms, respectively, N_e denotes the number desnity of electrons, N_{int} is the number density of neutral particles injected at the anode, $U_{i,w} = (e_{ch}T_e/m_i)^{1/2}$ is the ion acoustic speed, e_{ch} is the elementary charge of a particle, T_e is the electron temperature, m_i is the mass of a xenon ion, and L_{ch} is the channel length. This model assumes quasineutrality such that $N_e = N_i$ for the simulation. The spontaneous emission frequency ν is followed with the subscript (ml) for the deexcitation from Xe** to the metastable and radiative states and the subscript $(m0)$ for the deexcitation from the radiative state to the ground state. The reaction rate coefficients, k, are based on a series of curve-fit functions of electron temperature applied to data from Ref. 8 where the subscripts describe the reaction occurring. Here, the subscript D denotes direct ionization, s followed by m, r, or $**$ denotes stepwise ionization to one of the tracked excited states, and e followed by m, r, or $**$ denotes excitation to one of the tracked excited states. The two functional forms used to perform the curve fits are written as

$$k_1 = A\frac{\exp(B/T_e)}{\sqrt{T_e}}\sqrt{\frac{8e_{ch}T_e}{\pi m_e}}, \tag{3.20a}$$

$$k_2 = (A + BT_e)\frac{\exp C/T_e}{\sqrt{T_e}}\sqrt{\frac{8e_{ch}T_e}{\pi m_e}}, \tag{3.20b}$$

where A, B, and C are coefficients determined from the curve fit for each reaction rate and m_e is the mass of a xenon atom. The radiative and doubly excited excitation reaction rates use the k_1 form while all of the stepwise ionizations, direct ionization, and metastable excitation reaction rates use the k_2 form. A detailed presentation of these curve fits are presented in Appendix A.

3.3.5 Simulation Setup

As in the previous study, T_e is considered unknown and constant during the continuous prediction stage, but is corrected at every measurement update. The EKF uses the 0D ionization model shown in Eqs. (3.19a) - (3.19e) to propagate the particle number densities. Here, time-dependent

ion number density data, N_i, are calculated from the discharge current provided in Ref. 3 and used as the measurement data, $\tilde{\boldsymbol{y}}$. The beam utilization efficiency is treated the same as before.

The initialized value for the estimated ion number density, $\hat{N}_i$, is taken from the mean value of the measurement data. Using assumed ratios between particle number density sizes, the corresponding electron temperature can be calculated and all states given physically realistic initialization points. The initial variances, P_{ii} ($i = 1...6$), are set as $P_{ii} = (1 \times 10^{33}, 1 \times 10^{36}, 1 \times 10^{34}, 1 \times 10^{26}, 1 \times 10^{28}, 10)$. The mass flow rate is 5 mg/s and the average neutral velocity is assumed to be 400 m/s. The noise covariance matrices are initialized as $Q_{k=0} = 1 \times 10^{7}$ and $R_{k=0} = 1 \times 10^{15}$.

3.3.5.1 Physics-Constrained EKF Setup

The state vector of this model can now be written as $\boldsymbol{x} = \begin{bmatrix} N_i & N_n & N_m & N_r & N_{**} & T_e \end{bmatrix}^T$. This leads to the significantly more complex Jacobian matrix of

$$F(\boldsymbol{x}(t),t) = \begin{bmatrix} -U_i/L_{ch} + k_D N_n + k_{s,m} N_m + k_{s,r} N_r + k_{s,**} N_{**} & k_D N_i & k_{s,m} N_i \\ -k_D N_n - k_{e,m} N_n - k_{e,r} N_n, -k_{e,**} N_n & -U_n/L_{ch} - k_D N_i - k_{e,m} N_i - k_{e,r} N_i - k_{e,**} N_i & 0 \\ k_{e,m} N_n - k_{s,m} N_m & k_{e,m} N_i & -U_n/L_{ch} - k_{s,m} N_i \\ k_{e,r} N_n - k_{s,r} N_r & k_{e,r} N_i & 0 \\ k_{e,**} N_n - k_{s,**} N_{**} & k_{e,**} N_i & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} k_{s,r} N_i & k_{s,**} N_{**} & \partial RHS/\partial T_e \\ \nu_{m0} & 0 & \partial RHS/\partial T_e \\ 0 & \nu_{ml} & \partial RHS/\partial T_e \\ -k_{s,r} N_i - \nu_{m0} & \nu_{ml} & \partial RHS/\partial T_e \\ 0 & -k_{s,**} N_i - 2\nu_{ml} & \partial RHS/\partial T_e \\ 0 & 0 & 0 \end{bmatrix}, \tag{3.21}$$

where the $\partial RHS/\partial T_e$ is the partial derivative of each term involving a reaction rate coefficient, k with respect to T_e.

The mathematical constraint in Eq. (2.18) is used to maintain a positive covariance value at all times. A minimum covariance value is set to $P_{\min} = 0.1$ (see Eq. (3.11)) and a fixed variance for the unknown quantity, T_e, is fixed at $P_0 = 18$ for P_{66}^- (see Eq. (2.18)) to ensure that the state estimates satisfy the previously discussed mathematical constraints.

3.3.6 Results

Figures 3.11 and 3.12 show the true discharge current solution with the resulting estimated states in relation to their uncertainty bounds. These bounds are calculated assuming a 3σ standard deviation, where $\sigma = \sqrt{P_{ii}}$ $(i = 1...6)$. As these are the same cases as two of the four shown above, the impact of the excited states on the resulting estimates can be briefly mentioned. By including excited and metastable states in the model, there can be a greater distinction between fully and partially ionized particles in the model. Thus, the number density of neutrals can be seen to decrease slightly. The electron temperature estimates are largely unchanged by the increased fidelity of the model, indicating that using data-driven modeling can account for missing physics in lower-fidelity simulations.

The breathing mode shown in Fig. 3.11 is now solved using a beam ionization efficiency of 42% and reveals a coherent oscillation in all of the number densities, and an electron temperature with a time-averaged value of 18.9 eV compared to the 20.2 eV steady-state value estimated with the 3-state model using the same beam ionization efficiency (note that this is a different result than those presented in the previous section). Due to the highly oscillatory nature of this mode, the confidence region can be seen to fall into negative values for the ions, radiative state, and higher electronically excited state, more noticeably than in the 3-state model. While this could be considered unphysical, this is a result of the mathematics of the EKF rather than the physics. No conditions were placed on the variances to maintain strictly positive confidence bounds when subtracted from the estimate, though the variances themselves remain positive for all time. This could be added as an additional constraint to the EKF or the solutions could be presented with

only a 2σ uncertainty bound. Additionally, smaller noise covariances could reduce the size of the uncertainty bounds but may cause issue with the overall estimation.

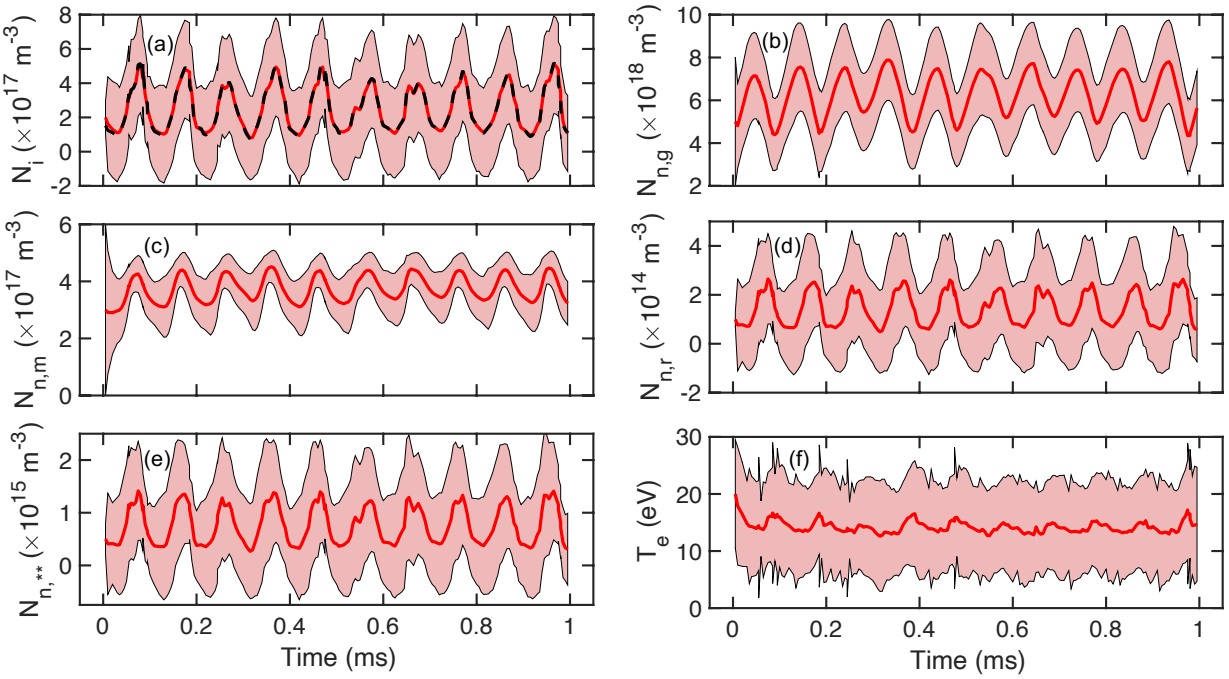

Figure 3.11: The breathing mode estimation solutions (red) shown with their 3σ confidence bounds (pink) for all state parameters in the system. The confidence bounds crossing zero are a result of the mathematics of the system without any positive-definite constraint in relation to the estimate. (a) The ion number density estimate and measurement (black dash). (b) Neutral (ground state) number density. (c) Metastable neutral state. (d) Radiative neutral state. (e) Electronically excited neutrals. (f) Electron temperature.

Figure 3.12 shows the stable discharge oscillation case from Ref. 3. As seen in the previous section, the random noise present in this operating mode are noticeable in all of the estimates after the initial transient period of the filter. Here, the electron temperature is estimated as 18.9 eV for a beam utilization efficiency of 57%. The significantly smaller oscillations in this data result in confidence bounds that do not cross zero, unlike in Fig. 3.11. As before, the neutral number density has a lower time-averaged value than the 3-state solution due to the inclusion of excited and metastable states.

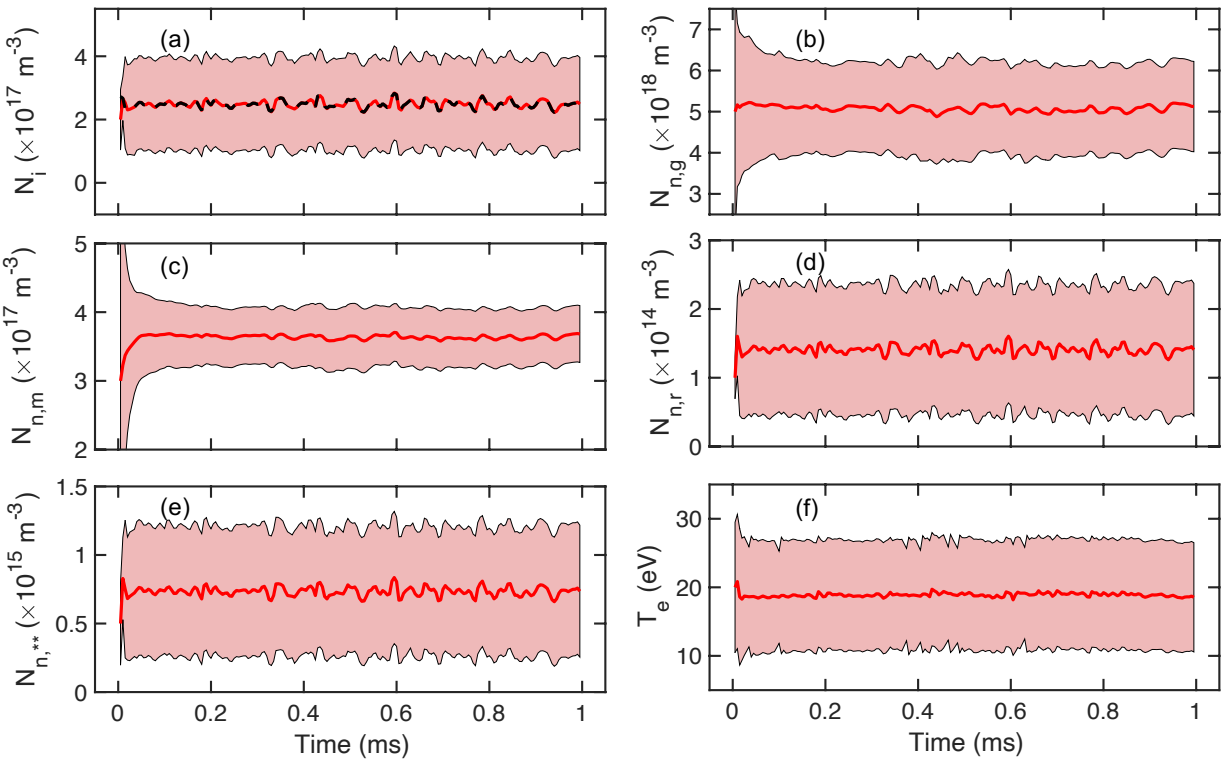

Figure 3.12: The steady-state estimation solutions (red) shown with their 3σ confidence bounds (pink) for all state parameters in the system. (a) The ion number density estimate and measurement (black dash). (b) Neutral (ground state) number density. (c) Metastable neutral state. (d) Radiative neutral state. (e) Electronically excited neutrals. (f) Electron temperature.

These Hall effect thruster test cases demonstrate the ability of the filter to estimate an unknown electron temperature in a physically consistent manner. This is a significant indication that an extended Kalman filter can be used for plasma applications where electron dynamics are of interest. Even though a limit has been found regarding how sparse of a measurement can be used to capture the complete dynamics of a signal, every solution has demonstrated that the time-averaged value of the filter estimates are physically consistent. Considering how little is known regarding electron dynamics and parameters such as electron mobility, even a result that shows an unknown electron dynamic varying in time without the precise high frequency oscillations of the dynamics would be a significant contribution to the research community.

4. ZERO-DIMENSIONAL PLASMA PROCESSING MODELS

For the great doesn't happen through impulse alone, and is a succession of little things that are brought together.

- Vincent Van Gogh, *The Hague, Oct. 1882*

Numerous research studies have developed zero-dimensional (0D) models to compare with experimental data and higher-dimensional models to demonstrate qualitative agreement of plasma processes that can provide useful, quick first-order approximations [38, 39, 40] These models employ the reaction rate coefficients for electron collisions, calculated by integrating collisional cross sections over an assumed energy distribution, as well as transport and diffusion into and out of the discharge region. Global models neglect the complexities that arise from spatial variations and encompass a large number of reactions. The results are not expected to yield accurate values of plasma parameters but rather to track overall trends and relations between different parameters at relatively low computational cost. The key challenge to these models is how best to simplify a system that can rely on spatial variation and dimensions. Furthermore, the input power supply, list of included reactions, and the shape of energy distributions can all play significant roles in simulation outcomes and must be carefully considered in these models.

4.1 Argon Global Model

Inductively coupled plasmas (ICPs) are low-pressure plasma sources that can operate as a steady-state or pulsed plasma dependent on the constant or modulated nature of the input power. Due to the scalability and simplicity of the setup of these sources, ICPs are widely studied for fundamental plasma phenomena. In this section, the capabilities of the PC-EKF are demonstrated with a ramped input power source for an argon ICP to estimate the absorbed input power. Then,

Part of this chapter is reprinted with permission from "Estimation of plasma properties using an extended Kalman filter with plasma global models" by C.M. Greve and K. Hara, 2022, *Journal of Physics D: Applied Physics*, 55, 255201, COPYRIGHT 2022 by IOP Publishing.

the PC-EKF is used to estimate reaction rate coefficients to study the limitations of the filter when tasked with estimating more hidden physics.

4.1.1 The Plasma Global Model

For any plasma discharge, tracking the ion (N_{Ar^+}), neutral (N_{Ar}), and single excited state (N_{Ar^*}) number density fluctuations can provide a wealth of information about the overall plasma. The particle balance equations are given by summing the reaction rates for different generation and loss processes of a given species. In particular, these processes are often a factor of a reaction rate coefficient, k, which reflects how likely a given chemical reaction, such as ionization, collisions, or excitation, will occur for a given electron temperature. These rate coefficients typically come from publicly available databases, though each database uses a unique method of calculating or measuring the rate coefficient of interest, leading to differences in coefficients between databases [121, 122]. With such non-standardized and imperfect calculation and extrapolation methods, inherent uncertainties exist within each rate coefficient due to measurement error or approximations in theoretical estimation.

For this study, a pure argon model described by 11 reactions as well as inflow and outflow of the ground state species is used, as shown in Table 4.1. The reaction rates are taken from various sources that assume a Maxwellian distribution of electrons [123, 124, 125]. The decision to use so few reactions was made to keep the both the physics-based model and PC-EKF simple for better understanding of the resulting estimates.

In Table 4.1, $N_{Ar,0}$ is the reference argon neutral number density based on the assumed operating gas temperature, the diffusion coefficient of excited state species to the wall (i.e., quenching) is given by

$$D_{[Ar^*]} = \frac{3}{16 N_{\text{tot}}} \frac{(2\pi k_B T_g / M_r)^{1/2}}{\pi \sigma_{12} \Omega_{12}}, \tag{4.1}$$

the ambipolar diffusion coefficient is considered for ion diffusion to the walls, given by

Table 4.1: A list of argon reactions included in the model and their corresponding rate coefficients where T_e is the electron temperature in eV.

Reaction		Type	Rate Coefficient	Reference
$e + Ar$	$\rightarrow Ar + e$	Elastic	$k_1 = 3.9 \times 10^{-7} \exp[-4.6/(T_e + 0.5)]$ cm^3/s	[122]
$e + Ar$	$\rightarrow Ar^* + e$	Excitation	$k_2 = 2.5 \times 10^{-9} T_e^{0.74} \exp(-11.6/T_e)$ cm^3/s	[39]
$e + Ar$	$\rightarrow Ar^+ + e + e$	Ionization	$k_3 = 2.3 \times 10^{-8} T_e^{0.68} \exp(-16/T_e)$ cm^3/s	[124]
$e + Ar^*$	$\rightarrow Ar + e$	Superelastic	$k_4 = 4.3 \times 10^{-10} T_e^{0.74}$ cm^3/s	[39]
$e + Ar^*$	$\rightarrow Ar^+ + e + e$	Multistep Ionization	$k_5 = 6.8 \times 10^{-9} T_e^{0.67} \exp(-4.4/T_e)$ cm^3/s	[124]
Ar^*	$\rightarrow Ar$	Radiation Trapping	$k_6 = 1 \times 10^5 (N_{Ar,0}/N_{Ar})$ s^{-1}	[126]
$e + Ar^+$	$\rightarrow Ar^*$	Radiative Recomb.	$k_7 = 4.3 \times 10^{-13} T_e^{-0.63}$ cm^3/s	[127]
$e + e + Ar^+$	$\rightarrow Ar^* + e$	Collisional Rad. Recomb.	$k_8 = 1.95 \times 10^{-27} T_e^{-4.5}$ cm^6/s	[128]
$Ar^* + Ar^*$	$\rightarrow Ar^+ + Ar + e$	Penning Ionization	$k_9 = 1.2 \times 10^{-9}$ cm^3/s	[129]
Ar^+ ($\rightarrow$ wall)	$\rightarrow Ar$	Ion Neutralization	$k_{10} = D_{[Ar^+]}/\Lambda^2$ s^{-1}	[130]
Ar^* ($\rightarrow$ wall)	$\rightarrow Ar$	Metastable Quenching	$k_{11} = D_{[Ar^*]}/\Lambda^2$ s^{-1}	[130]

$$D_{[Ar^+]} = \left(1 + \frac{T_e}{T_i}\right) \mu_0 \frac{N_0}{N_{\text{tot}}} \frac{k_B T_i}{e}, \tag{4.2}$$

and Λ is the diffusion length calculated as $1/\Lambda^2 = (\pi/L)^2 + (2.405/R)^2$ assuming a cylindrical plasma discharge. Here, N_{tot} is the total particle number density, T_g is the gas temperature in Kelvin, T_i is the ion temperature in Kelvin, M_r is the reduced mass between species j and the neutral species, σ_{12} is the Lennard-Jones radius, Ω_{12} is the non-deal correction factor (taken as 1 for this work), μ_0 is ion mobility at standard temperature and pressure, N_0 is the gas density at standard temperature and pressure, e is the elementary charge, and L and R are the length and radius of the simulated geometry, respectively.

Based on the reactions included in this model as described in Table 4.1, the volume-averaged equations can be written as

$$\begin{aligned}\frac{\partial N_{Ar}}{\partial t} = -k_2 N_e N_{ar} - K_3 N_e N_{ar} + k_4 N_e N_{Ar,m} + k_6 N_{Ar,m} + k_9 N_{Ar,m}^2 + k_{10} N_{Ar^+} \\ + k_{11} N_{Ar,m} + \Gamma_{in,Ar} - \Gamma_{out,Ar},\end{aligned} \tag{4.3a}$$

$$\frac{dN_{Ar,m}}{dt} = -k_4 N_e N_{Ar,m} - k_5 N_e N_{Ar,m} - k_6 N_{Ar,m} - 2K_9 N_{Ar,m}^2 - k_{11} N_{Ar,m} + k_2 N_e N_{Ar} + k_7 N_e N_{Ar^+} + k_8 N_e^2 N_{Ar^+}, \tag{4.3b}$$

$$\frac{dN_{Ar^+}}{dt} = -k_7 N_e N_{Ar^+} - k_8 N_e^2 N_{Ar^+} - k_{10} N_{Ar^+} + k_3 N_e N_{Ar} + k_5 N_e N_{Ar,m} + k_9 N_{Ar,m}^2, \tag{4.3c}$$

where the variable Γ, the particle flux, tracks the inflow and outflow of the neutral species and N_e is the number density of electrons which under quasineutral conditions is equal to the number density of ions, i.e., $N_e \approx N_i$. These volume-averaged rate equations can be used to simulate the dynamic plasma behavior and determine the state of the plasma at any time based on a driving electron temperature equation. The electron temperature evolution can be solved from a power balance equation as,

$$\frac{d}{dt}\left(\frac{3}{2} n_e k_B T_e\right) = \frac{P_{\text{abs}}}{V} - \sum_{\text{j=elastic}} R_j \Delta\varepsilon_j - \sum_{\text{j=inelastic}} R_j \Delta\varepsilon_j, \tag{4.4}$$

where n_e is the electron number density, k_B is the Boltzmann constant, T_e is the electron temperature in Kelvin, P_{abs} is the absorbed input power, V is the system volume, R_j are the reaction rates for each process, and $\Delta\varepsilon_j$ are the associated electron energy loss terms. Electron energy losses to the surface are not directly accounted for in this model as the estimation results are expected to contain information about this physical process.

4.2 Simulation Setup

A self-verification study is performed using a time-dependent reference solution generated by providing a power ramp and subsequent constant input power to the electron temperature equation. The resulting ion flux at every point in time can be calculated for use as the measurement data in the PC-EKF. The plasma global model, Eqs. (4.3) and (4.4), combined with Table 4.1 constitute the physics-based model in Eq. (2.1) within the EKF framework. The reference solution is generated with a power ramp from 0 W to 100 W for 200 μs and then held at 100 W for the remainder of the simulation. the pressure is set as 50 mTorr, the flowrate is 300 sccm, and the GEC reference

cell chamber geometry has a radius of 40 cm and an axial height of 3 cm. The gas temperature is set equal to the ion temperature at 300 K. The physics-based model is solved using a time step of 100 ns, and a data acquisition frequency, i.e., the frequency of the discrete Kalman update, of 100 kHz.

The incoming measurement ($\tilde{y}$) is the ion flux density, calculated as

$$\Gamma_{Ar^+} = N_{Ar^+}\sqrt{\frac{eT_e}{m_{Ar}}}, \tag{4.5}$$

where e is the elementary charge, T_e is the electron temperature in electron-volts, and m_{Ar} is the argon mass assuming Bohm condition.

4.2.1 Physics-Constrained EKF Setup

The state vector for this system contains the three particle species number densities, the electron temperature, and the unknown state of interest. For the following studies, the unknown state will include the absorbed electron power, one reaction rate coefficient, or two reaction rate coefficients. Using the case of the unknown absorbed electron power, the state vector can be constructed as $\boldsymbol{x} = [N_{Ar^+}\ N_{Ar}\ N_{Ar^*}\quad T_e\ P_{abs}]^T$.

This state vector, combined with the equations given in Sec. 4.2.2, leads to the following Jacobian matrix for the covariance propagation equation:

$$
F(\boldsymbol{x}(t),t) = \begin{bmatrix}
-k_7N_e - k_8N_e^2 - k_{10} & k_3N_e & k_5N_e + 2k_9N_{Ar,m} & \partial RHS/\partial T_e \\
k_{11} & -k_2N_e - k_3N_e & k_6 + 2k_9N_{Ar,m} + k_{10} & \partial RHS/\partial T_e \\
k_7N_e + k_8N_e^2 & k_2N_e & -k_4N_e - k_5N_e - k_6 - 4k_9N_{Ar,m} - k_{11} & \partial RHS/\partial T_e \\
0 & 1/N_e(2/(3e_{ch}))(k_1N_e\epsilon_1e_{ch} - k_2N_e\epsilon_2e_{ch} - k_3N_e\epsilon_3e_{ch}) & 1/N_e(2/(3e_{ch}))(-k_4N_e\epsilon_4e_{ch} - k_5N_e\epsilon_5e_{ch}) & 1/N_e(2/(3e_{ch}))(-\partial K_{tot}/\partial T_e - dN_e/dt) \\
0 & 0 & 0 & 0
\end{bmatrix},
\tag{4.6}
$$

where $\partial RHS/\partial T_e$ refers to the time derivative of the right hand side of each equation where time derivatives of each reaction rate coefficient are taken, ϵ are the threshold energies of each reaction, and $\partial K_{tot}/\partial T_e$ is the partial derivative of the first five reaction rates in Table 4.1, i.e., $k_1 - k_5$, with respect to electron temperature.

The initial variances are $\hat{P}_{mm} = (1 \times 10^{30}, 1 \times 10^{34}, 1 \times 10^{30}, 10, 50)$, where $m = 1, 2, 3, 4, 5$, while all off-diagonal terms (covariances) are initially set to zero. The noise covariance values are initialized as $Q = 1 \times 10^7$ and $R = 1 \times 10^{15}$, but are immediately updated by the physics-based constraints using values of $P_{min} = 0.01$ and $P_0 = 50$ for this test case.

4.2.2 Self-Verification with Unknown Absorbed Electron Power

The incoming measurement, $\tilde{y} = \Gamma_{Ar^+}$, is a function of two state variables, N_{Ar^+} and T_e, as shown in Eq. (4.5). This test demonstrates how the PC-EKF can use a convolution of more than one state as a single measurement signal and correctly decompose the parts of the measurement

signal. Note that the xenon Hall effect thruster plasma global model in Sec. 3.1 considers the ion density as the measurement signal, a signal that corresponds to an exact state of interest in the system. The key intrigue in this new setup is that there are infinite possibilities for the plasma system that would return the measured ion flux density, but only one that estimates the complete plasma dynamics. This is of particular interest looking towards future plasma physics applications where measurements such as discharge current are commonly available [3], yet comprise numerous plasma parameters that could convolute the application of the PC-EKF.

Figure 4.1 depicts the estimated ion flux, ion number density, electron temperature, and absorbed electron power states (solid red line) compared to the reference solutions (dashed black line) for the case of an unknown ramped input power. The initial oscillations shown in the electron temperature estimate, Fig. 4.1(c), are caused by the transient nature of the state estimates while the PC-EKF searches for the true dynamics of the system. The electron temperature is estimated at a reasonable order of magnitude and is shown to decrease with an increase in absorbed power. Simultaneously, the ion number density is shown to increase as the input power increases. Having access to the full state of the reference solution, the exact ion flux solution can be plotted against the estimate in Fig. 4.1(a). The PC-EKF provides an accurate estimation of the unique plasma properties based on the ion flux measurement dataset and physics-based model without *a priori* knowledge of the electron power input. Furthermore, the only information given to the PC-EKF regarding the relation of the two states observed through the measurement signal is comprised in the observation vector H, where the partial differential equations of the ion flux with respect to each state is given.

A similar study is performed assuming that one of the reaction rate coefficients is unknown in the system while the absorbed electron input power is known, with the same ion flux measurement signal. Despite reaction rate coefficients being considered hidden parameters in plasma chemistry models, the PC-EKF is able to obtain a reasonable estimate for the coefficient of interest, as demonstrated in Fig. 4.2. As the solution is nearly identical to the absorbed electron input power results, these studies are not discussed in greater detail. Instead, the more interesting case of estimating

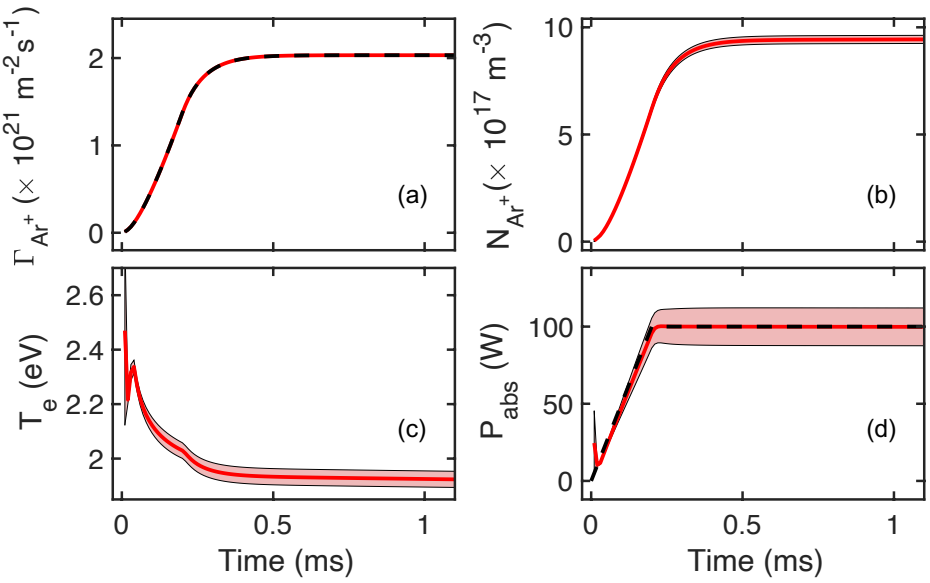

Figure 4.1: The EKF estimates (red solid line) using an ion flux reference dataset (black dashed line) as the measurement to estimate the unknown absorbed electron power. (a) The measured and estimated argon ion flux. (b) The estimated ion number density with 3σ uncertainty bounds (pink shade). (c) The estimated electron temperature with 3σ uncertainty bounds. (d) The reference and estimated input power time history with 3σ uncertainty bounds.

more than one unknown, simultaneously, is discussed below.

4.2.3 Simultaneous Estimation of Two Unknown Reaction Rate Coefficients

Having demonstrated that the filter can estimate relevant plasma parameters individually, the PC-EKF is used to estimate two unknowns simultaneously to determine the effectiveness of the filter under increased uncertainty in the model. The additional unknown state is added to the end of the state vector with all equations and matrices being modified to follow suit. The process noise covariance, Q, becomes a 2-by-2 matrix whose diagonal entries relate to each unknown state such that the constraint equations from Sec. 2.2 are solved separately for each entry in Q. This requires two separate P_{min} values which are set as $P_{min,1} = 1 \times 10^{-32}$ and $P_{min,2} = 1 \times 10^{-34}$. The measurement noise covariance are set using a P_0 condition of $P_{0,1} = 1 \times 10^{-28}$ and $P_{0,2} = 1 \times 10^{-32}$.

For the following results, the excitation (k_2 in Table 4.1) and multistep ionization (k_5) rate coefficients are used as the unknown states, indicating that no dynamics related to these two rate

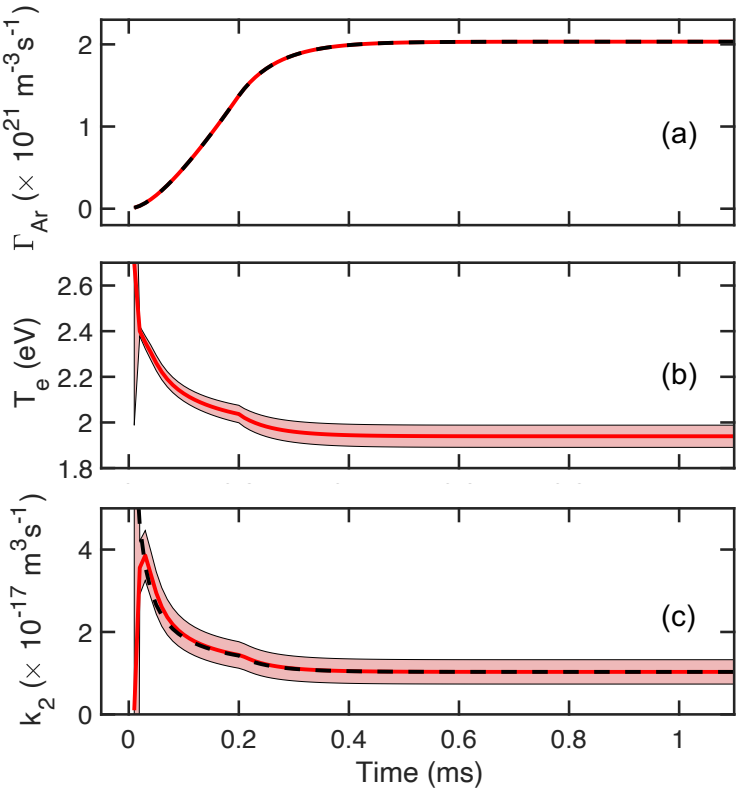

Figure 4.2: The EKF estimates (red solid line) using an ion flux reference dataset (black dashed line) as the measurement to estimate the excitation reaction rate coefficient, k_2. (a) The measured and estimated argon ion flux. (b) The estimated electron temperature with 3σ uncertainty bounds. (c) The reference and estimated excitation reaction rate coefficient with 3σ uncertainty bounds.

coefficients are known by the physics-based global model; they are assumed piecewise constant during the continuous propagation phase. The filter is run using the same geometry and operating conditions as in Sec. 4.2.2, with the initialized covariance matrix written as $\hat{P}_{mm} = (1 \times 10^{30}, 1 \times 10^{34}, 1 \times 10^{30}, 10, 1 \times 10^{-32}, 1 \times 10^{-30})$ for $m = 1, 2, 3, 4, 5, 6$ with all other terms equal to zero. The ion flux density, as shown in Eq. (4.5), serves as the incoming measurement signal.

Before analyzing the results, it is pertinent to note that for the given model under the specified conditions, the excitation reaction rate coefficient is one of the most significant processes in the model, rivaled only by the ionization process. In comparison, all other reactions in the model occur at such a low rate their rate coefficients need not be perfectly matched to recreate the ion flux measurement.

Figure 4.3 shows the resulting estimates from the PC-EKF under the conditions provided above. Due to the greater significance of the excitation reaction to the overall state of this global plasma model, it can be seen that the excitation rate (k_2) is estimated more accurately than the multistep ionization (k_5). After analyzing the data, the multistep ionization reaction rate coefficient was found to not be updated by the incoming experimental data at any time. Subsequent studies demonstrated that the PC-EKF gives preference to the dominant reactions, properly estimating those unknown reaction rates compared to less significant states when multiple unknowns exists in the system, explaining why only one unknown was estimated well by the filter. Despite the poor multistep ionization reaction rate estimate, the electron temperature is shown to be estimated at nearly the same steady-state value as in Fig. 4.1. Thus, despite the multistep ionization rate coefficient never being updated by the PC-EKF, the overall solution of the filter can be trusted.

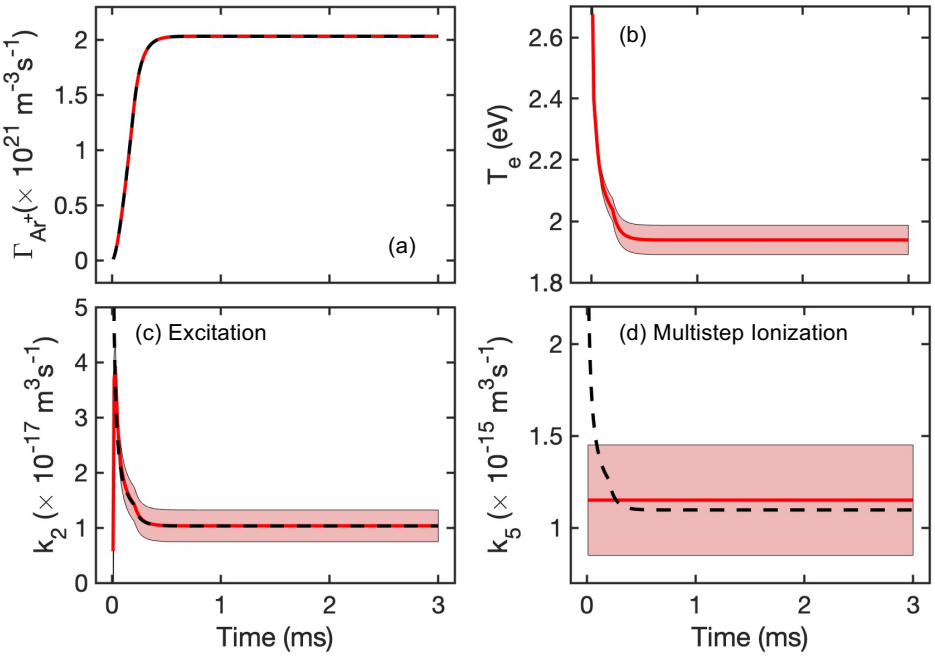

Figure 4.3: Estimation of the (a) ion flux, (b) electron temperature, (c) excitation rate coefficient, and (d) multistep ionization rate coefficient for a global argon ICP model using the physics-constrained EKF. As in Fig. 4.1, the 3σ uncertainty bounds (pink shade) are overlaid with the estimate (red solid line) and reference solution (black dashed line).

These observations provide an important consideration for future usage of the EKF model for plasma systems with multiple unknowns. If the filter can recognize the significant reactions for a given operating condition, the EKF technique could be used to develop reduced-order models that ignore or approximate less critical processes. These less sensitive states and parameters could be identified in *real time* by tracking the Kalman gain terms of the filter, which indicate how large of an update is performed on each individual state, to understand which processes are not necessary for certain operating conditions or dynamics of interest. Following this notion, further tests were run to determine if it were possible for the PC-EKF to update more than one unknown state or if it would always update the more significant state. Figure 4.4 shows a similar case where both the elastic and superelastic collision rate coefficients are estimated. It is immediately evident that the PC-EKF updates both unknown states, as well as all other states in the state vector without violating any physical laws. The increased transient time for the filter to find a steady-state solution, or even match the reference dynamics, is likely indicative of the increased challenge and uncertainty regarding two equally significant unknown parameters in the system.

Notably, the EKF *failed* for the particular case when both the excitation (k_2) and ionization (k_3) rate coefficients are unknown in the physics-based model, meaning that they are not propagated during the continuous prediction phase. These reaction rate coefficients are the two most critical processes in the plasma generation for this particular simple model. Without knowledge of these two processes, the filter is unable to find a viable solution due to the inability of the physics-based model to propagate forward in time despite a variety of initial conditions for the unknown reaction rates. It is unlikely that future studies will include physics-based models missing such fundamental or critical processes. Nevertheless, this theory indicates that cases where the EKF cannot run may be useful to determine the required fidelity of the physics-based model.

4.2.4 Pulsed Mode Experimental Study with Unknown Driving Absorbed Electron Power

Operating a plasma in pulsed-mode can alleviate some undesirable effects in plasma processing applications, such as sputtering, wafer erosion, and silicon notching due to insulator charging [131], as pulsing the plasma creates new control variables [1, 132]. Global models using an

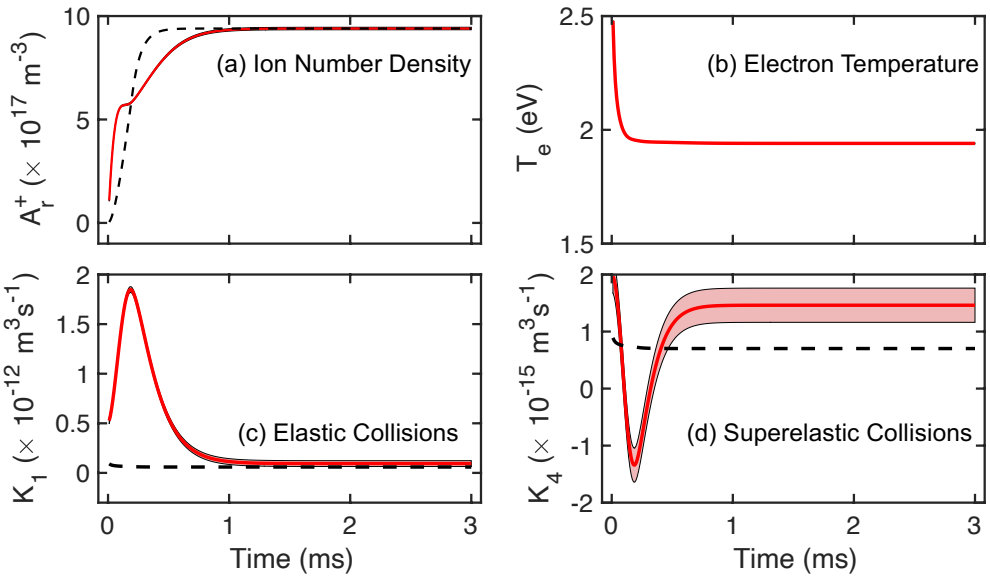

Figure 4.4: Estimation of the (a) ion number density, (b) electron temperature, (c) elastic collisions rate coefficient, and (d) superelastic collisions rate coefficient for a global argon ICP model using the physics-constrained EKF. As in Fig. 4.1, the 3σ uncertainty bounds (pink shade) are overlaid with the estimate (red solid line) and reference solution (black dashed line).

assumed absorbed power value report large discrepancies, up to factors of five, between the calculated and measured plasma density [133]. Hence, while qualitative agreement with experiments can be obtained with these global models, some plasma properties may not be in quantitative agreement. This indicates that the input to the existing models may be incomplete, such as the power absorption to the electrons. This parameter is *a priori* not known, as the RF power from the circuit is used not only for electron heating, but also for volume and surface chemistry, radiation, heating of materials and components, and ion acceleration [134]. Data-driven modeling can be useful for estimating such *hidden* or inaccessible properties due to the inclusion of real-time measurement data and the ability to generalize unknown states to estimate the effects of more than one process in a single variable.

The physics based model used for this study is identical to the argon global model developed for the previous tests. The state vector is taken as $\boldsymbol{x} = \begin{bmatrix} N_{Ar^+} & N_{Ar} & N_{Ar^*} & T_e & P_{\text{abs}} \end{bmatrix}$ where the unknown state is again the electron absorbed input power, P_{abs}. As with all previous cases, the

electron absorbed input power is only corrected during every measurement update and is assumed constant during the physics-based propagation. The initial state variances are $\hat{P}_{mm} = (1\times10^{30}, 1\times10^{34}, 1\times10^{30}, 10, 5)$ for $m = 1, 2, 3, 4, 5$ with off-diagonal terms set to zero while the measurement noise covariances are initialized as $Q = 1 \times 10^{7}$ and $R = 1 \times 10^{15}$. The measurement signal frequency is 50 MHz for all cases. The physics-constrained variables are initialized as $P_{min} = 0.01$ and $P_0 = 1$ and the geometry is that of the GEC reference cell discussed in the previous section.

This PC-EKF includes an additional constraint such that the absorbed power estimate is constrained to zero if the measurement updates result in a negative value. This constraint ensures a consistent physics-based model as a negative absorbed power causes the physics-based model to diverge. The previous constraints on the PC-EKF only require the variances of the unknown state to meet a positive-definite condition to enable estimates of states that may physically be negative. Because a negative input power is detrimental to this particular model, a specific additional constraint could be included. Such constraints on the updated solutions may be required in future studies to ensure physical relevance of more complex plasma models.

The measurement signal is the time-dependent argon electron number density for two cycles of the pulsed mode as it steadily operates from Ref. 1 which is treated as the ion number density due to a quasineutral assumption for the plasma. Reference 1 does not specify all operating conditions of the experiment, so a flowrate of 10 sccm is assumed in the following results. The experimental data from Ref. 1 are extracted using an online digitizer. This extraction results in slight numerical inaccuracies in the reference data, but still allows the demonstration of the PC-EKF.

As mentioned during the setup of the argon ICP model, the electron temperature equation does not take the energy loss of electron losses to the wall into account. Though this decision is made to keep the equations relatively simple for the PC-EKF, the resulting absorbed input power estimate can be seen to account for these physical processes not explicitly included in Eq. (4.4). The estimates provided by the EKF are considered the effective absorbed input power as opposed to the true absorbed input power and are expected to be lower than the true value due to the missing physical processes.

4.2.4.1 Effects of Power Inputs on Plasma Estimates

Figure 4.5 shows the estimates for ion number density, electron power absorption, and electron temperature for three peak power cases, 165 W, 190 W, and 290 W, using the same frequency pulse (10 kHz) and duty cycle (30%) based on the reported experimental results in Ref. 1. Note that peak power refers to the RF power supplied to the induction coil, which is different from the power absorbed by the electrons.

The resulting estimates of the absorbed electron input power are presented in Figs. 4.5(a2), (b2), and (c2). With an increase in peak power, there is an evident increase in the estimated absorbed electron power. For all cases, the initial electron (ion) density is low at the time when the measurement data begins. As ionization occurs, the electron density increases and causes an increase in absorbed power as the plasma is generated. During the *on* phase, the power absorption reaches some maximum value and begins to decay, likely due to the physical notion that once the plasma is formed and the ions are accelerated to the walls, the input power is used more for ion acceleration than electron heating. Additionally, the higher RF peak power values lead to faster saturation during the *on* phase. It is evident that the PC-EKF is able to estimate the pulsed nature of the absorbed electron power input using a single measurement datastream without *a priori* knowledge of the pulsed nature of the discharge. The PC-EKF is shown to recognize that the absorbed power needs to be near-zero within one measurement timestep of the experimental data.

Figures. 4.5(a1), (b1), and (c1), demonstrate that the plasma density estimates are in good agreement with the measurements. Note that the decay rate during the afterglow phase when the power is turned off is slightly underpredicted by the PC-EKF as shown by the ion density being estimated slightly higher than the measurement data. This underprediction of the rate of decay is likely caused by the electron temperature estimates, as shown in Figs. 4.5(a3), (b3), and (c3). The electron temperature first increases to approximately 3 eV, consistent with the increase in absorbed electron power. As the plasma density continues to increase, the electron temperature is nearly constant until the absorbed electron power begins to decrease in the *on* phase. The electron temperature decreases to approximately 2.5 eV still during the *on* phase as the absorbed electron power

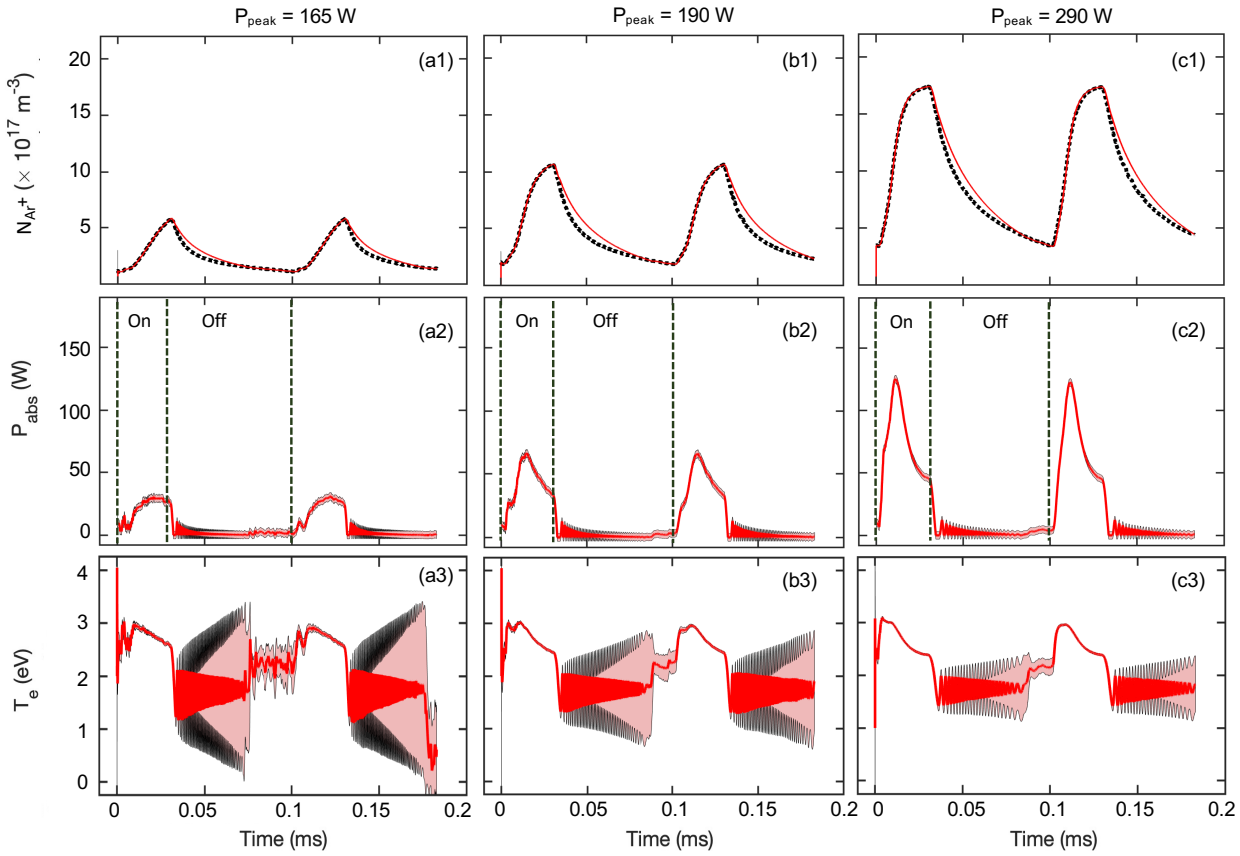

Figure 4.5: Comparison of absorbed electron power over the course of two pulses in the ICP based on Ref. 1 with a duty cycle of 30% and a pulse repetition frequency of 10 kHz. The peak powers are (a) 165 W, (b) 190 W, and (c) 290 W, respectively. (1) The original ion number density measurements (black dashed line) with the resulting estimates (solid red line), (2) the estimated absorbed power trends with the 3σ uncertainty bounds (pink shade), and (3) the electron temperature estimates with their 3σ uncertainty bounds.

is transferred to accelerating ions. Immediately after the electron power absorption is estimated to be zero by the PC-EKF, the electron temperature estimate reduces to about 1.5 eV. It is notable that the similar *off* phase electron temperature values lead to similar decay rates in the plasma for the same duty cycle. The electron temperature oscillates during the *off* phase in the present EKF model as a numerical artifact caused by the additional positive constraint placed on the PC-EKF update. The power *off* phase shows a noticeable increase in the uncertainty bounds for the electron temperature. Such an increase in uncertainty illustrates the difficulty the filter has in finding the true zero-power operation mode.

Analyzing the absorbed power estimates reveals a discrepancy between the peak powers provided in the referenced work and the estimated power absorbed by the plasma. The absorbed power estimated by the PC-EKF is approximately only one-third of the reported peak power from Ref. 1. Other computational studies have attributed low absorbed power rates to capacitive coupling in the plasma, ion acceleration under the coil, resistive power loss, and other processes [134]. Some studies comparing experimental and computational results note that correction is needed to achieve quantitative agreement in the models to ensure accurate physical dynamics are being described by the model [133]. Other studies cite power dissipation caused by well-known sources including exciting or ionizing neutral particles by electron collisions and the transfer of kinetic energy to the wall from electrons and ions [135, 136].

4.2.4.2 Effects of the Duty Cycle on State Estimates

The effect of duty cycle on the PC-EKF estimates for a peak power of 300 W and a pulse frequency of 10 kHz are presented at 10%, 30%, 50%, and 70% in Ref. 1. All four cases are used with the PC-EKF in this section.

Figure 4.6 continues to demonstrate good estimation performed by the PC-EKF. The power *off* periods are oscillatory, as when studying the effect of peak RF power, caused by the filter searching for a non-zero solution. These oscillations can be seen to damp over time as the *off* periods increase. In the second pulse of Fig. 4.6(a3), the filter shows a drop in estimated electron temperature toward the end of the *off* phase, which indicates an absorbed power estimate that finds a steady, zero-value solution. In other cases, the filter finds a slightly positive value for the absorbed electron power or does not have sufficient time to relax to a non-oscillatory solution.

As the duty cycle increases, the filter fails to capture the afterglow phase as accurately as with the shorter duty cycles, possibly due to the inability to find a non-oscillatory absorbed power estimate before the next pulse begins. Note that the plasma density at the exact moment when the RF power is turned on is lower for the shorter duty cycle cases due to the operating condition of the background gas for each different case. This results in a larger absorbed electron power for the plasma to be generated. This can be seen from Figs. 4.6(a2), b(2), (c2), and (d2) where the

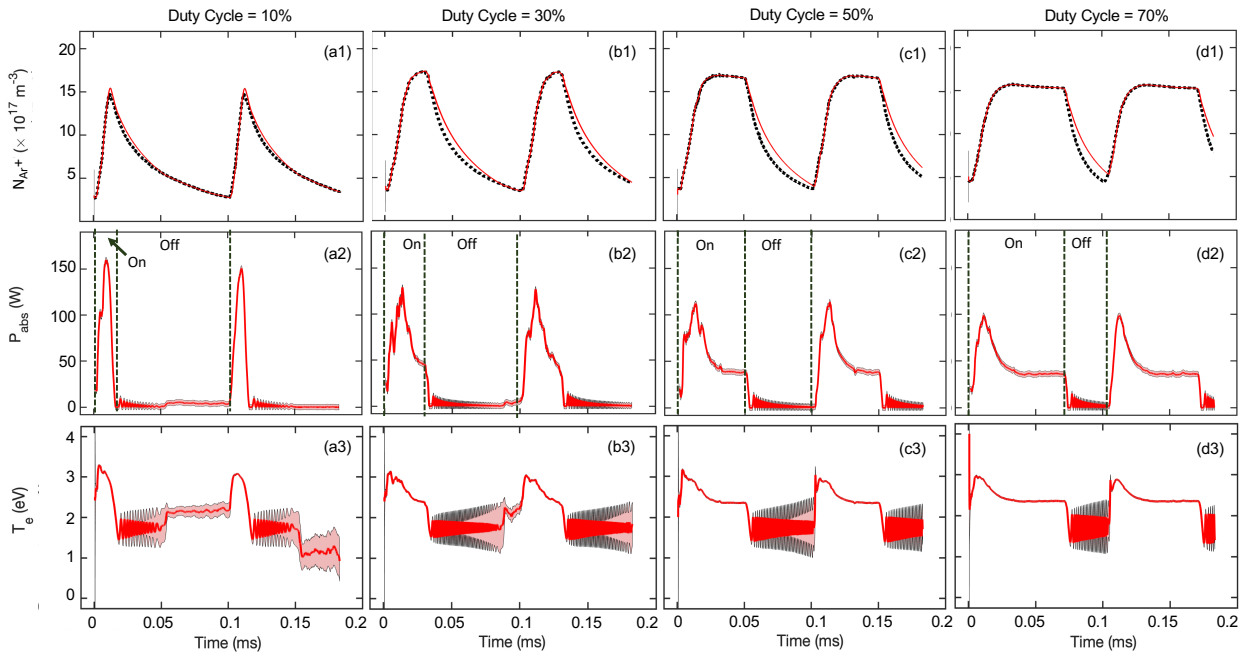

Figure 4.6: The effect of the duty cycle on the pulsed plasma dynamics. Here, duty cycles of (a) 10%, (b) 50%, and (c) 70% are studied based on experimental data from Ref. 1 with a peak power of 300 W and a pulse frequency of 10 kHz. (1) The ion number density measurements (black dashed line) compared to the estimated values (red solid line), (2) the corresponding absorbed electron power estimates with their 3σ uncertainty bounds (pink shade), and (3) the electron temperature estimates and their 3σ uncertainty bounds.

maximum power absorption decreases as the duty cycle increases. The maximum electron power absorption is estimated to be 160 W for the 10% duty cycle case while it only reaches 100 W for the 70% duty cycle case. A higher ion density value during the *off* phase requires less absorbed electron power to setup the plasma, hence why the peaks drop in magnitude as the duty cycle increases but the steady plateau absorbed power once the ions are being accelerated is approximately constant between test cases.

4.2.4.3 Effects of the Pulse Frequency on State Estimates

The PC-EKF is applied to study the effect of varying pulse frequency while the peak power is held constant at 300 W with a 50% duty cycle [1] for three different frequencies.

Figure 4.7 once again shows that the filter estimates the pulsed mode using the time-dependent experimental data of the plasma density. Shorter pulses return less accurate estimates of the plasma

density during the power *off* phase, seen in Figs. 4.7(a1), (b1), and (c1), and consistent with the results shown in Figs. 4.5 and 4.6. The peak absorbed electron power in Figs. 4.7(a2), (b2), and (c2) is seen to decrease with a shortened pulse length due to the increase in initial ion number density between cases, although the steady absorbed power value for the sustained plasma condition remains the same. For the case of Fig. 4.7(a3), the power estimate damping towards a zero condition leads to a further decrease in the electron temperature while the other two cases, Figs. 4.7(b3) and (c3), do not exhibit this trend. Also note that as the pulse frequency increases, the uncertainty bounds of the electron temperature do not grow as large during the *off* phase of the pulse. This is largely due to the shorter time between non-zero power estimates, allowing the filter to maintain better confidence in the solution over time.

4.3 Argon-Oxygen Global Model

The use of gas mixtures in plasma processing applications can provide added stability across a range of operating conditions. Thus, the PC-EKF model is applied to an 80/20 argon-oxygen plasma global model using rate coefficients from literature [137, 138, 139, 140] and experimental data from Ref. 2. This testcase will demonstrate the PC-EKF with a significantly more complex model than has been tested up to this point in the dissertation. The physics-based model uses the argon global model developed in Sec. 4.2.2 with twenty-five oxygen reactions and three oxygen-argon reactions. These new reactions introduce the neutral state, O, vibrationally excited state of diatomic oxygen, O_2(v), electronically excited state of diatomic oxygen, O_2^*, diatomic oxygen ion, O_2^+, and monoatomic negative oxygen ion, O^-, to the physics-based model and state vector. Note that the oxygen species and reactions are not complete [141] as the focus of this study is to demonstrate the capabilities of PC-EKF. Increasing the fidelity in the physics-based model is reserved for future work.

Tests estimating various reaction rate coefficients and the absorbed electron power input were performed using the argon-oxygen model to ensure that the added complexity did not noticeably hinder the performance of the PC-EKF but are not presented as part of this dissertation. Instead, this section investigates the effects of supplying different combinations of unknown states and mea-

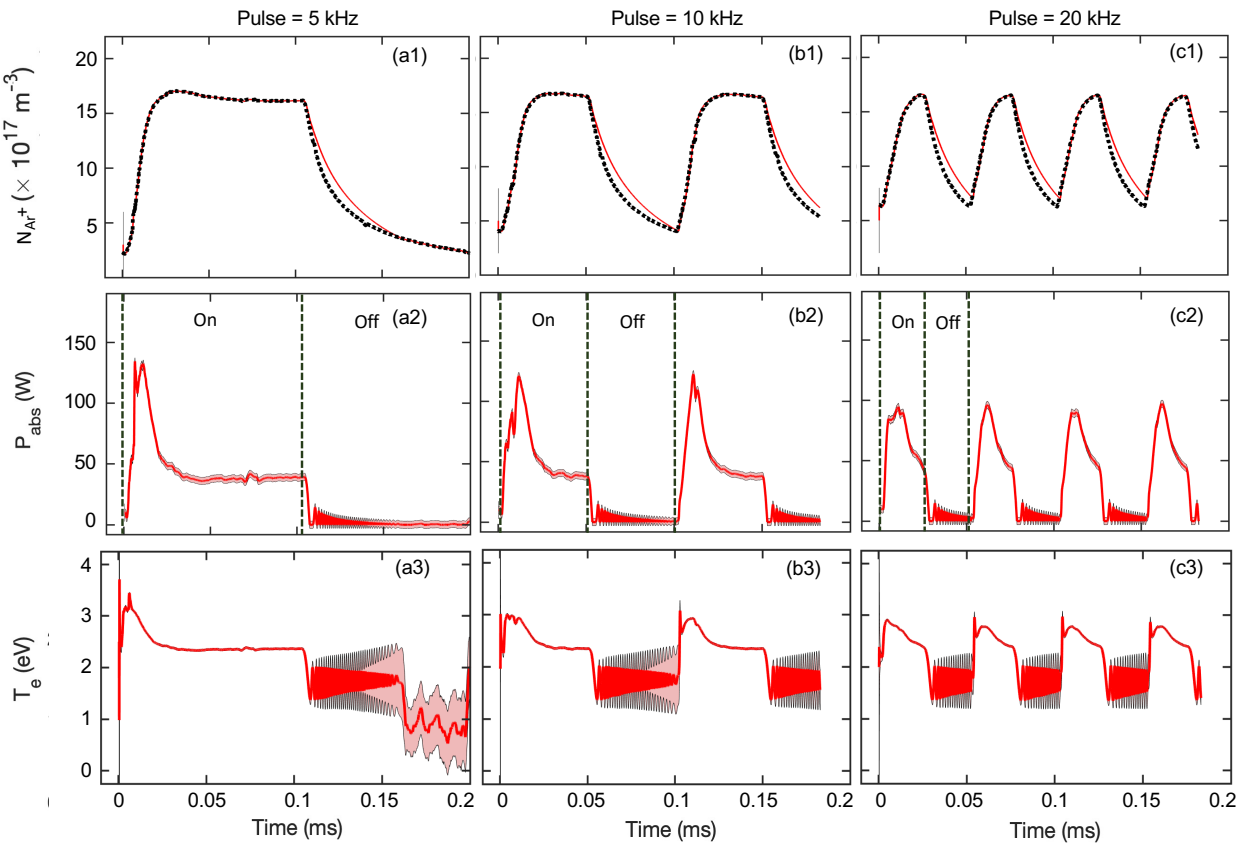

Figure 4.7: Effects of the pulse frequency on the pulsed plasma dynamics. Here, frequencies of (a) 5 kHz, (b) 10 kHz, and (c) 20 kHz are studied based on experimental data from Ref. 1 with a constant peak power of 300 W and a 50% duty cycle. (1) The ion number density measurements (black dashed line) compared to the estimated values (red solid line), (2) the corresponding absorbed electron power estimates with their 3σ uncertainty bounds (pink shade), and (3) the electron temperature estimates with their 3σ uncertainty bounds.

surement signals to the PC-EKF. Due to the mixed gas nature of this test, the PC-EKF is essentially tasked with estimating two streams of information simultaneously, one for the argon portion of the plasma and one for the oxygen portion. This adds new complexity to the physics-based model and reveals an example of when including a second measurement signal can significantly improve the system estimates.

4.3.1 Simulation Setup

The argon-oxygen model is constructed using the argon reactions given in Table 4.1 with the addition of the oxygen and oxygen-argon reactions provided in Table 4.2. For the neutral oxygen

wall recombination reactions, the sticking coefficient is set as $\beta = 0.05$.

Table 4.2: A list of oxygen and argon-oxygen reactions included in the model and their corresponding rate coefficients where T_g is the gas temperature in K, D is the wall diffusion coefficient as described in Sec. 4.2.2, and β is the sticking coefficient. [*] uses detailed balance to determine the reaction rate coefficient.

Reaction		Type	Rate Coefficients	Reference
$e + O_2$	$\rightarrow O_2 + e$	Elastic	$k_{12} = 4.79 \times 10^{-8} T_e^{0.5}$ cm^3/s	[137]
$e + O_2$	$\rightarrow O + O + e$	Dissociation	$k_{13} = 6.86 \times 10^{-9} \exp(-6.29/T_e)$ cm^3/s	[138]
$e + O_2$	$\rightarrow O + O + e$	Dissociation	$k_{14} = 3.49 \times 10^{-9} \exp(-5.92/T_e)$ cm^3/s	[138]
$e + O_2$	$\rightarrow O + O^-$	Dissociative Attachment	$k_{15} = 1.07 \times 10^{-9} T_e^{-1.39} \exp(-6.26/T_e)$ cm^3/s	[138]
$e + O_2$	$\rightarrow O_2(v) + e$	Vibrational Excitation	$k_{16} = 2.80 \times 10^{-9} \exp(-3.72/T_e)$ cm^3/s	[139]
$e + O_2$	$\rightarrow O_2(v) + e$	Vibrational Excitation	$k_{17} = 1.28 \times 10^{-9} \exp(-3.67/T_e)$ cm^3/s	[139]
$e + O_2$	$\rightarrow O_2^* + e$	Electronic Excitation	$k_{18} = 1.37 \times 10^{-9} \exp(-2.14/T_e)$ cm^3/s	[138]
$e + O_2$	$\rightarrow O_2^+ + e + e$	Ionization	$k_{19} = 2.34 \times 10^{-9} T_e^{1.03} \exp(-12.3/T_e)$ cm^3/s	[138]
$e + O_2(v)$	$\rightarrow O_2 + e$	Superelastic	$k_{20} = 2.80 \times 10^{-9} \exp(-3.53/T_e)$ cm^3/s	[*]
$e + O_2(v)$	$\rightarrow O_2^+ + e + e$	Ionization	$k_{21} = 2.34 \times 10^{-9} T_e^{1.03} \exp(-12.11/T_e)$ cm^3/s	[139]
$e + O_2(v)$	$\rightarrow O + O + e$	Dissociation	$k_{22} = 6.86 \times 10^{-9} \exp(-6.10/T_e)$ cm^3/s	[138]
$e + O_2(v)$	$\rightarrow O + O + e$	Dissociation	$k_{23} = 3.49 \times 10^{-9} \exp(-5.73/T_e)$ cm^3/s	[138]
$e + O_2(v)$	$\rightarrow O + O^-$	Dissociative Attachment	$k_{24} = 1.07 \times 10^{-9} T_e^{-1.39} \exp(-6.26/T_e)$ cm^3/s	[140]
$e + O_2^*$	$\rightarrow O_2 + e$	Superelastic	$k_{25} = 2.06 \times 10^{-9} \exp(-1.162/T_e)$ cm^3/s	[138]
$e + O_2^*$	$\rightarrow O_2^+ + e + e$	Ionization	$k_{26} = 2.34 \times 10^{-9} T_e^{1.03} \exp(-11.32/T_e)$ cm^3/s	[138]
$e + O_2^*$	$\rightarrow O + O + e$	Dissociation	$k_{27} = 6.86 \times 10^{-9} \exp(-5.31/T_e)$ cm^3/s	[140]
$e + O_2^*$	$\rightarrow O + O + e$	Dissociation	$k_{28} = 3.49 \times 10^{-9} \exp(-4.94/T_e)$ cm^3/s	[139]
$e + O_2^*$	$\rightarrow O + O^-$	Dissociative Attachment	$k_{29} = 1.07 \times 10^{-9} T_e^{-1.39} \exp(-6.26/T_e)$ cm^3/s	[140]
$O^- + Ar^+$	$\rightarrow O + Ar$	Ion-ion Neutralization	$k_{30} = 3 \times 10^{-7}$ cm^3/s	[142]
$O^- + O_2^+$	$\rightarrow O_2 + O$	Ion-ion Neutralization	$k_{31} = 2.6 \times 10^{-8} (300/T_g)^{0.44}$ cm^3/s	[138]
$e + O_2^+$	$\rightarrow O + O$	Dissociative Recombination	$k_{32} = 2.2 \times 10^{-8} T_e^{-0.5}$ cm^3/s	[137]
$Ar^+ + O_2$	$\rightarrow O_2^+ + Ar$	Charge Exchange	$k_{33} = 4.9 \times 10^{-11} (300/T_g)^{0.78}$ cm^3/s	[138]
$Ar^* + O_2$	$\rightarrow O + O + Ar$	Dissociative Quenching	$k_{34} = 1.0 \times 10^{-10}$ cm^3/s	[138]
O_2^+ ($\rightarrow$ wall)	$\rightarrow O_2$	Ion Neutralization	$k_{35} = D_{[O_2^+]}/\Lambda^2$ s^{-1}	[138]
O_2^* ($\rightarrow$ wall)	$\rightarrow O_2$	Metastable Quenching	$k_{36} = D_{[O_2^*]}/\Lambda^2$ s^{-1}	[138]
$O_2(v)$ ($\rightarrow$ wall)	$\rightarrow O_2$	Vibrational Quenching	$k_{37} = D_{[O_2(v)]}/\Lambda^2$ s^{-1}	[138]
O ($\rightarrow$ wall)	$\rightarrow (\beta/2)O_2$ $\rightarrow (1-\beta)O$	Recombination	$k_{38} = D_{[O]}/\Lambda^2$ s^{-1}	[138]

The same GEC reference cell geometry is used from the previous section and similar operating conditions are considered, though the inflow and outflow quantities are split between the neutral and metastable argon and diatomic oxygen species for this testcase. Ion flux data for the total flux, and argon ion, diatomic oxygen, and monoatomic negative oxygen species reported experimentally in Ref. 2 are used as the measurement signal, for continuous operation at 30 mTorr and 100 W.

4.3.1.1 Physics-Constrained EKF Setup

The PC-EKF is provided a state vector with either 11 or 12 entries, dependent on the study, to include all species in the model, the electron temperature, and unknown state(s), i.e., $\boldsymbol{x} = [N_{Ar^+}\ N_{Ar}\ N_{Ar^*}\ N_{O_2}\ N_{O_2(v)}\ N_{O_2^*}\ N_{O_2^+}\ N_O\ N_{O^-}\ T_e\ \xi]^T$. Additional unknown states can be included by appending them to the end of the state vector. Initialized quantities for the various covariances are set as follows: the model error covariance is $Q = 1 \times 10^{12}$, the measurement error covariance is $R = 1 \times 10^{15}$, and the covariance matrix diagonal entries are $\hat{P}_{mm} = (1 \times 10^{34}, 1 \times 10^{36}, 1 \times 10^{34}, 1 \times 10^{40}, 1 \times 10^{36}, 1 \times 10^{36}, 1 \times 10^{32}, 1 \times 10^{40}, 1 \times 10^{32}, 15, 5)$ for $m = 1, 2, 3, ..., 11$.

4.3.2 Results

The initial studies revealed that the PC-EKF was unable to find a physically consistent solution for the low pressure cases below 10 mTorr. Based on recent PIC/MCC simulations [143], it was determined that the rate of ions diffusing to the wall may be influenced by the plasma profile, whose nonuniformity is not captured by the volume-averaged 0D model. Further literature review suggests that ion diffusion to the wall depends on the area of the sheath [39], total gas density [144], and wall temperature [145], which change with the variation of total gas pressure and electron temperature. Reference 146 also demonstrates that different internal pressures can exist between the argon and oxygen populations by studying how plasma conditions change with different partial pressures of O_2. Thus, an effective wall diffusion factor is added to the physics-based plasma global model to account for the inhomogeneity of the plasma flow in the ICP source. This is done by creating a constant multiplied by the assumed wall diffusion (ion neutralization) reaction rate equation assuming a uniform plasma profile such that

$$k_{11} = \alpha_{Ar^+} \frac{D_{[Ar^+]}}{\Lambda^2}, \tag{4.7}$$

and

$$k_{35} = \alpha_{O_2^+} \frac{D_{[O_2^+]}}{\Lambda^2}, \tag{4.8}$$

where α_{Ar^+} and $\alpha_{O_2^+}$ are the effective ion diffusion factor for Ar^+ and O_2^+, respectively. The PC-EKF is written to estimate the effective ion wall diffusion factors, α_{Ar^+} and $\alpha_{O_2^+}$, as the unknown state(s).

4.3.2.1 Effects of the Measurement Signal

Initial studies attempted to only use one measurement signal to inform the state of the entire system, as well as only studying a single effect wall diffusion coefficient. As can be seen in Fig. 4.8, which compares the solutions if only one ion flux is used or both. When only the argon ion flux is provided to the EKF, as seen by the teal solid line in Fig. 4.8, the EKF will properly capture that measurement at the cost of all other estimates. The oxygen ion flux can be seen to severely underestimate the true amount of ion flux over time, Fig. **??**(a). This is caused by the EKF not having sufficient information to correctly update all states of interest in the model due to the limited interaction of the two species. Conversely, using both ion flux measurement signals with only the argon effective wall diffusion term yields a better overall estimate where both fluxes are closer to the measurement signal, but neither flux is seen to match the measurement signals. The addition of the second ion flux measurement is also shown to increase the transient time it takes the filter to settle into a steady-state operation compared to the one measurement signal. It is decided that the EKF needs to have greater control over the physics-based model to improve the estimates and match both measurement signals, so the effective wall diffusion for oxygen is added as an unknown.

4.3.2.2 Effects of Multiple Unknowns

Having demonstrated the improved results of using more than one measurement signal, this physics-based model is now used to demonstrate how adding more controllable parameters to the PC-EKF can improve the overall simulation results. The necessity of using an ion diffusion factor for both argon and diatomic oxygen introduces a test case where two unknowns of equal significance to the system exist, but are not of such high significance that the model cannot propagate forward in time without them, unlike the ionization and excitation reaction rate coefficients in

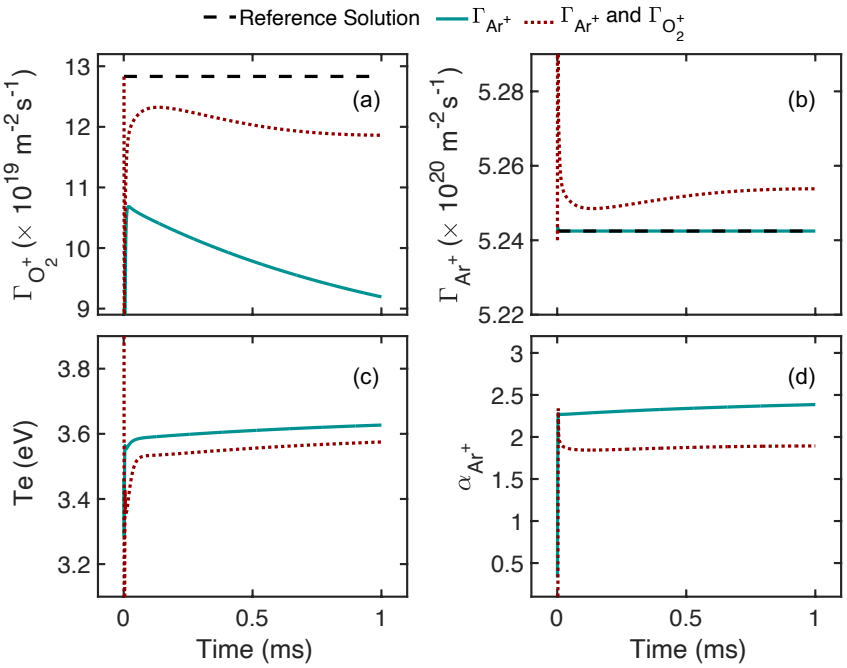

Figure 4.8: Effect of using one measurement signal (teal line) compared to two measurement signals (red dot) for the effecting argon wall diffusion coefficient in the argon/oxygen global model. The EKF estimates are shown for (a) diatomic oxygen flux, (b) argon flux, (c) electron temperature, and (d) the effective argon wall diffusion term. The true flux reference solutions are shown in black dash.

Table 4.3: List of Argon-Oxygen EKF Cases

Case	Measurement	Unknown(s)	Fixed Values
I	$\Gamma_{Ar^+}, \Gamma_{O_2^+}$	α_{Ar^+}	$\alpha_{O_2^+} = 1$
II	$\Gamma_{Ar^+}, \Gamma_{O_2^+}$	$\alpha_{O_2^+}$	$\alpha_{Ar^+} = 1$
III	$\Gamma_{Ar^+}, \Gamma_{O_2^+}$	$\alpha_{Ar^+}, \alpha_{O_2^+}$	-

Sec. **??**. Thus, the ability of the filter to estimate one, the other, or both unknown ion diffusion factors is presented in this section as demonstrated by the different test cases in Table 4.3. Each case uses the ion flux measurement signals of both argon and oxygen to inform the filter, but the estimation varies from searching only for the argon or oxygen effective wall diffusion terms to both coefficients simultaneously.

Figure 4.9 compares the results of the cases listed in Table 4.3. Case I, i.e., when α_{Ar^+} is

allowed to be updated by the PC-EKF but $\alpha_{O_2^+} = 1$ is assumed to be constant, demonstrates reasonable estimates for the electron temperature, Fig. 4.9(e) and the argon ion wall diffusion factor, α_{Ar^+}, Fig. 4.9(c). As a consequence of assuming that $\alpha_{O_2^+} = 1$, the oxygen ion flux, $\Gamma_{O_2^+}$ is noticeably underestimated due to an artificially increased diffusion to the wall. Stated alternatively, no *controllable* oxygen process exists for the PC-EKF that enables the state estimation to agree with the oxygen ion flux data. This is due to the lack of correlation between argon and oxygen in this model with only three reactions providing information on the interaction between species. The underprediction of the oxygen flux causes the filter to find a higher operating condition of the argon ion flux to create a physically-consistent solution, as shown in Figs. 4.9(a) and (b).

Case II, i.e., when $\alpha_{Ar^+} = 1$ is constant and $\alpha_{O_2^+}$ is treated as a time-varying unknown that can be updated by the PC-EKF, reproduces the oxygen ion flux, $\Gamma_{O_2^+}$, to high levels of accuracy as shown in Fig. 4.9(b). The argon ion flux is largely overestimated as shown in Fig. 4.9(a). The estimated electron temperature, Fig. 4.9(e), lies outside of the uncertainty bounds for Case III, considered the most correct solution. This indicates a significant divergence from the true state of the measured system as the uncertainty bounds constitute 99.7% of the values within a normal distribution and the Case II results lie outside of such bounds. By comparing Cases I and II, it is evident that being able to estimate the dominant processes in such a mixed-gas system continues to be significant to correctly estimating the overall state of the plasma. Despite the Case I results poorly matching the ion flux values for both species, the state estimates for all states of interest are significantly closer to those of Case III, as shown in Figs. 4.9(c-e). This is because argon ion density is significantly larger than that of the oxygen species and therefore carries greater influence over the state of the system.

Case III, i.e., when both α_{Ar^+} and $\alpha_{O_2^+}$ are considered unknown, estimates both ion fluxes in good agreement with the measurement data, Figs. 4.9(a) and (b). This result indicates that allowing both species of plasma to be controlled by the PC-EKF can improve the overall estimate of the system. Note that the electron temperature is approximately 3.5 eV, as shown in Fig. 3(e). This is higher than the pure argon case in Sec. 4.2.2 and consistent with other literature that note

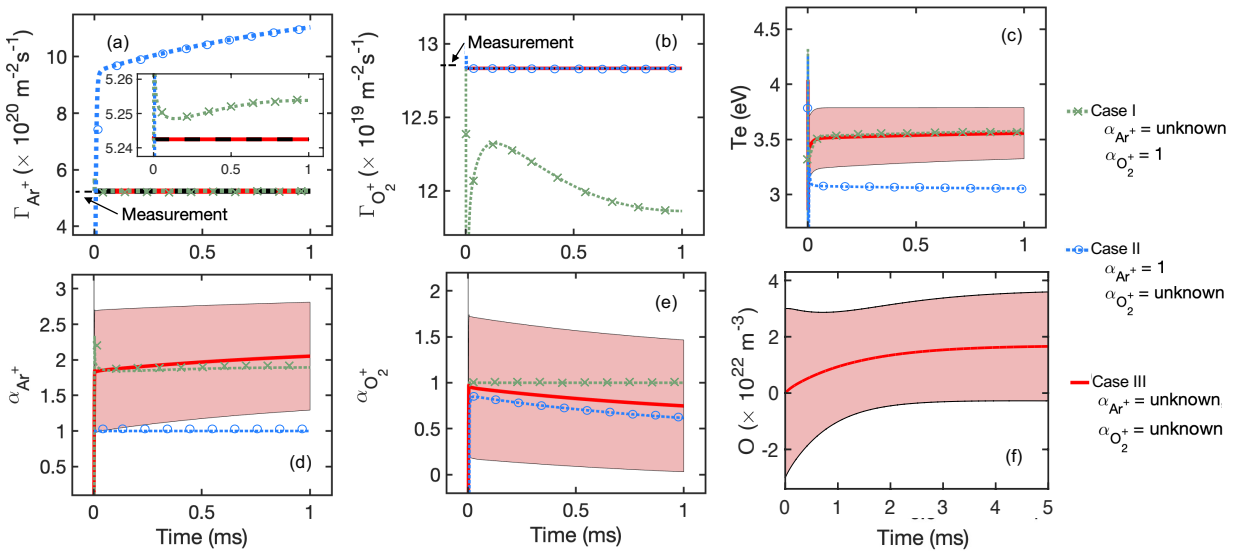

Figure 4.9: Estimating the effective ion wall diffusion factors, α_{Ar^+} and $\alpha_{O_2^+}$, in the argon-oxygen global model for the cases shown in Table 4.3. (a) Estimated argon ion flux density. (b) Estimated oxygen ion flux density. Measurement data are taken from Ref. 2. (c) Estimated argon ion wall diffusion factor. (d) Estimated oxygen ion wall diffusion factor. (e) Estimated electron temperature. (f) The atomic oxygen number density estimate shown for 5 ms to illustrate that the EKF reaches a steady state. Note that the results in (a-e) focus on the initial 1 ms transient for a clear comparison between the three cases.

mixed-gas plasmas require higher power levels to maintain an ionized plasma [147].

The results for each case presented in this section reach a steady state solution if the simulation is run for 5 ms as shown in Fig. 4.9(f), due to the slow relaxation time for oxygen processes [148], but are not shown for every plasma parameter to better present the transient portion of the filtering process. The atomic oxygen number density is chosen for this figure to demonstrate how the steady-state value is consistent with trends seen in pure-oxygen configurations in literature [148]. Case III shows that the effective ion wall diffusion factor reaches 2.203 and 0.624 for argon and oxygen ions, respectively, at steady state, validating the theory that an oxygen wall diffusion coefficient equal to one was too large and resulted in too much depletion of the oxygen plasma. The difference in wall diffusion factors between species indicates that the spatial profile of the argon and oxygen ion densities are not identical within the chamber. Reference 149 shows the existence of a variation in the plasma density profile between argon and oxygen ions using a computational

model for different operating conditions of argon-oxygen plasmas than those presented in this section.

Over these studies, the extended Kalman filter is demonstrated to reach different estimate profiles dependent on the number of measurements and number of unknowns in the system. Put another way, the controllability given to the EKF regarding the physics of the system can greatly improve or greatly hinder the resulting estimates. This study demonstrates that using two ion flux measurement signals and allowing the effective wall diffusion factors of both diatomic oxygen and argon ions to be estimated yields the most accurate estimates based on the available experimental data.

5. ONE-DIMENSIONAL TEST CASES

...don't become a slave to the model...take a model and study it, for otherwise your inspiration won't take on material form.

- Vincent Van Gogh, *The Hague, Nov. 1882*

Having verified the EKF with zero-dimensional models, gaining understanding of its abilities and limitations, the filter is extended for application with one-dimensional models. Significantly less literature exists regarding the use of an EKF with one-dimensional systems, though a handful of oceanography papers were found as part of this study. In the oceanography studies, the EKF is typically applied to a coarse spatial grid that tracks a low number of states in each cell [150]. Considering the thought experiment of a domain with ten cells tracking three states in each cell, it is immediately evident that such an approach can quickly lead to filters tracking tens and hundreds of states, increasing computational cost. This can be problematic for plasma simulations that typically have hundreds of cells and numerous states of interest within each individual cell.

Before approaching the challenge of maintaining low computational costs as the systems being studied become more complex, an understanding of how to utilize the EKF in one spatial dimension is necessary. One-dimensional plasma models have been widely used to study plasma phenomena across all manner of applications. These models provide insight to how plasma densities and velocities are distributed across the computational domain. The collisional nature of plasma inherently leads to location-determined physics occurring, unable to be captured by global, zero-dimensional models. While plasma dynamics are inherently multi-dimensional in nature, the bulk of the phenomena of interest can be tracked using a single spatial dimension. For HETs, this is often the axial direction, using the assumption of azimuthal symmetry and providing contrived values to account for radial differences if required. A number of one-dimensional models have been developed in the plasma physics community to study device-scale physics in Hall effect thrusters, carbon arc discharges, and plasma distributions in inductively and capacitively coupled

plasmas.

For this section, three one-dimensional test cases are presented. First a linear advection propagation scheme is used to estimate the initial condition of the domain, recreating work done in Ref. 150. Then, a heat conduction problem is created to analyze the ability of the filter to estimate diffusion coefficients in different regions of the domain. Lastly, initial applications and challenges of the state estimation technique to a Hall effect thruster model are discussed.

5.1 Linear Advection

Advection is the transport of a substance or quantity by the bulk motion of a fluid, carrying the properties of the fluid with it. This motion is described by a partial differential equation that governs the motion of a conserved scalar field as it is advected by a known velocity vector field.

One particular oceanography study first demonstrated the use of the EKF with the reconstruction of a sinusoidal advection model Ref. 150. A true solution is developed for a simple wave propagation study with periodic boundary conditions. The EKF is constructed using a first-order upwind linear advection propagation scheme and provided a constant, zero-value initial condition for the domain. Two measurement signals, measuring the values of two different cells within the domain, are taken from the true solution to inform the state update of the EKF method. Using only these two measurement locations, the EKF is shown to reconstruct the initial domain within a few time steps. Due to the clarity with which the original study was presented, said study is recreated to gain familiarity with the one dimensional EKF application. To ensure that similar results are obtained to the study performed in Ref. 150, the original EKF will be used for this chapter, without the constraints previously developed as part of this dissertation. Thus, only constant values for Q and σ_R will be discussed in this chapter.

5.1.1 Simulation Setup

Propagation of a domain can be solved exactly to generate a measurement solution for a code-to-code verification study. This work studies the pure advection of y given a constant velocity c.

The initial condition is set using a sinusoidal distribution of $w_0 = \sin(2\pi x/\lambda)$ where $\lambda = L/2$ is the wavelength equal to half the domain and x is the cell location. The exact solution is solved assuming a constant propagation speed of $c = 1$ resulting in the exact solution of

$$w = \sin(2\pi(x - t)/\lambda) \tag{5.1}$$

where $x = n\Delta x$ is the position, n is the cell number, $\Delta x = 0.25$ is the cell width, and t is the current time, for 60 time units. The domain consists of 100 cells, $x \in [0, 30]$, and the solution is solved on the cell centers. The system uses $\Delta t = 0.06$ to solve the true solution over the course of four oscillations.

Upwind methods are one of a number of types of numerical discretization schemes for solving hyperbolic partial differential equations. The name comes from the theory where spatial difference are skewed in the direction from which the flow originates. The first-order approximation in both time and space creates a low-fidelity propagation solution that can lose information about the original state of the system over time. In particular, the partial differential equation of advection can be written as

$$w_t = -cw_x, \tag{5.2}$$

which can be solved exactly over a period of four oscillations. These propagation schemes rely on a stability condition based on the idea that the time step of the system must be smaller than the time taken for the wave to travel across a cell width. This stability condition is commonly called the Courant-Friedrichs-Lewy (CFL) stability criterion. This stability condition can be written as

$$\frac{c\Delta t}{\Delta x} \leq 1. \tag{5.3}$$

A first-order upwind propagation scheme is used to propagate every cell in the domain as

$$\frac{w_n^{k+1} - w_n^k}{\Delta t} = -c\frac{w_n - w_{n-1}}{\Delta x}, \tag{5.4}$$

where k indicates the time step and n the cell index.

5.1.2 EKF Setup

Unlike all previous applications of the EKF, here the system is attempting to estimate the initial condition based on known propagation equations. To accomplish this task, the covariance matrix becomes critical for informing the first few updates. To this end, the covariance matrix is set based on a correlation of cells in the nearby region using a user-defined scaling factor. Each cell is given a variance of one along the diagonal of the covariance matrix initially, while all cross-covariances are given a correlation based on the function

$$P_{ij} = \exp(-(r/r_e)^2), \tag{5.5}$$

where r is the location with respect to the reference point, Δx, and r_e is the user-defined folding scale, the width of influence for a particular point. The wider the region, the more Gaussian the cross variances are in shape. The effect of this parameter is shown in Fig. 5.1 where difference scale lengths, r_e, are plotted together for the initial covariance values surrounding the two measurement points. From this figure, it can be seen how a larger folding scale will affect more nearby states, possibly enabling a shorter transient time before the filter is able to recover the original state of the system.

The EKF is provided measurement data at $x = 5$ and $x = 20$ every 5 continuous model time steps, $\Delta t_m = 0.6$ units. Thus, as the system propagates, the EKF receives more information about the initial waveform, populating the initial condition as the system advects. Even though the measurements are perfect in this particular problem, the EKF is told to assume some error covariance for both measurement points, with the second measurement being given twice as high an error variance as the first. The states provided to the EKF include the solution for every cell, i.e., $\hat{x} = [w_1, w_2, w_3, ..., w_n]$. Thus, the Jacobian equations can be written as

$$\frac{\partial w_n^{k+1}}{\partial w_n^k} = \frac{-c}{\Delta x}, \tag{5.6a}$$

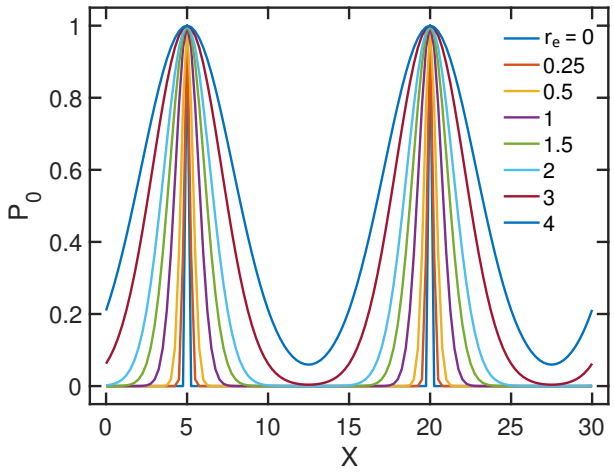

Figure 5.1: Example of the initialized covariance values across the domain for different r_e folding scale values. The presented results represent the summed values for the covariances related to both measurement locations to depict how the Kalman gain would be structured in each case.

$$\frac{\partial w_n^{k+1}}{\partial w_{n-1}^k} = \frac{c}{\Delta x}, \tag{5.6b}$$

for the central portion of the domain. The boundary conditions supplement what would be values w_0 and w_{101} with the cell on the opposite side of the domain due to the assumed periodic structure. As noted at the beginning of the chapter, the process noise covariance and measurement noise covariance are taken as user-defined constants in an effort to better match the literature reference rather than applying the physics-informed method developed earlier in this dissertation. A study on the effects of Q and σ_R will be performed as part of this test case, so no particular values are presented here. The r_e folding scale is set as $r_e = 3$ for all solutions in this section.

5.1.3 Reconstructing the Initial Domain

The original literature, Ref. 150, that studied this case presents results at 1, 4, and 11 time steps to show how the estimate improves over time. As the chosen EKF setup for this work is not perfectly identical to the published study, a comparison of results will not be made. Instead, a presentation of the results from this study is given. For this first study, the process noise covariance is set as $Q = 0.01$ and the measurement noise is considered $\sigma_{R,1} = 0.1$ and $\sigma_{R,2} = 0.2$ for the two measurements at $x = 5$ and $x = 20$, respectively.

The solution and corresponding covariance are presented after the first EKF update in Fig. 5.2(a). Immediately evident is the difference in amplitude height for each of the two measurement locations. This is a direct result of the difference in assumed error variance for the two measurement points while the width reflects the cross correlation between off-diagonal terms of the covariance matrix as demonstrated in Fig. 5.1. Examining the covariance solutions of the diagonal terms of the matrix, the variances of each state, it is evident that the errors are reduced in the near region of the measurements.

Figure 5.2(b) shows the same set of solutions after five update steps. Now that the initial wave structure has propagated a noticeable amount, the estimates are seen to approach the form and amplitude of the reference solution. The errors can be plotted to demonstrate a continued to decrease as the filter gains additional insight into the true solution, i.e., as more measurement data become available. Because of the propagation direction from left to right based on the upwind advection scheme used in this study, i.e., $c > 0$, the measurement information has propagated downstream. Notably, the errors have also propagated downstream. Thus, the errors downstream of the measurement location will be smaller than those upstream.

Figure 5.2(c) shows the solution state after 20 measurement updates. Plotting the errors reveals that they are regularly oscillating between certain minimum values immediately after the Kalman filter update and maximum values immediately before the update. Due to the selected propagation scheme and model noise, the errors will always increase during the continuous propagation phase while they are reduced during the measurement update phase. As the model errors have decreased, the Kalman gain amplitudes have decreased, indicating that the EKF is beginning to trust the model more heavily than the incoming measurements.

Figure 5.2(d) shows the solution after 55 update steps, which corresponds to one oscillation. This solution is nearly identical to Fig. 5.2(c) in terms of the visual closeness of the solution to the reference solution as well as the shape and size of the error bounds. While first-order advection schemes are known to be dissipative, the EKF estimates do not show such an effect. This is caused by the continued information supplied to the EKF as measurement data continue to arrive

after the initial domain has been constructed. The minimal reduction in amplitude height over the course of this simulation gives evidence to how the EKF can be used to generate more reliable and physically consistent solutions with low fidelity models in addition to estimating unknown states and parameters.

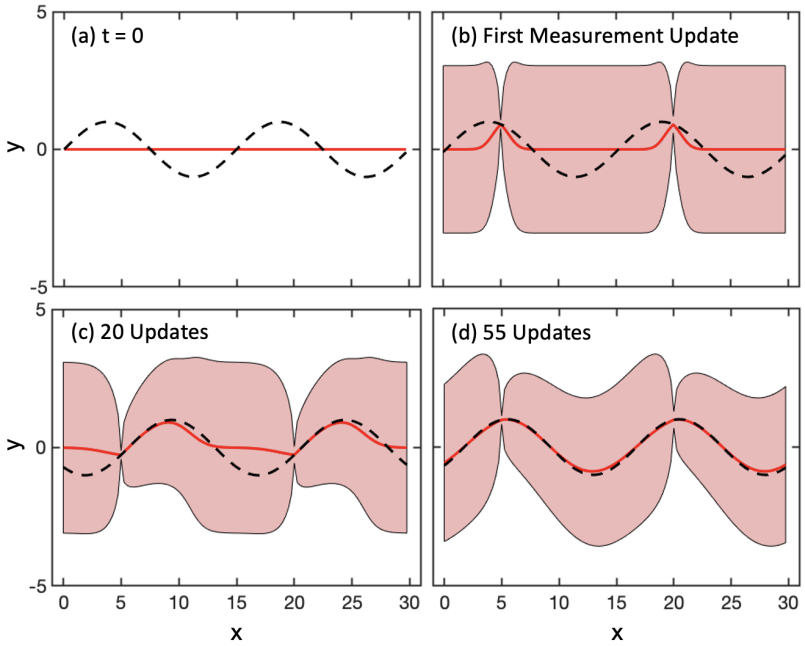

Figure 5.2: The EKF estimates (red solid line) using the reference dataset (black dashed line) at two locations as the measurement to reconstruct the initial domain. (a) The initial condition of the EKF and reference solution. (b) The estimated domain after 5 update steps with 3σ uncertainty bounds (pink shade). (c) The estimated domain after 20 update steps with 3σ uncertainty bounds. (d) The estimated domain after 55 update steps with 3σ uncertainty bounds.

Having determined that the EKF can, indeed, be used to reconstruct a sinusoidal domain using a linear advection propagation scheme, further studies were performed to analyze the effect of the process and measurement noises on the final estimates. The process noise measurement accounts for inaccuracies in the physics-based model as it propagates forward in time. Figure 5.3 presents four different cases of process noise values for the same measurement noise of 0.05 and 0.1 at locations $x = 5$ and $x = 20$, respectively. The case of no process noise measurement, $Q = 0$, is shown in Fig. 5.3(a). Due to the lack of error assumed in the model, the solution is seen to not

reach the full oscillation height of the true solution. In addition, the uncertainty bounds are hardly noticeable, and, at times, imaginary values. The error in the estimates can be calculated using the L2 norm which is written as

$$|\boldsymbol{x}| = \sqrt{\sum_{k=1}^{n} |x_l|^2},$$

for any vector $\boldsymbol{x}$. This solution yields a L2 norm of 3.48821, mainly due to its inability to replicate the full oscillation amplitude. Figure 5.3(b) shows the case where $Q = 0.1$ to indicate some slight error in the propagation model. Even with such a low value, the estimate can be seen to significantly improve. The L2 norm becomes 0.729884 and the uncertainty bounds clearly encapsulate the true solution across the entire domain. Further increasing $Q = 1$ enlarges the uncertainty bounds in Fig. 5.3(c). Here, the bounds are sufficiently large to indicate that the errors are smallest at the measurement locations and increase the further away from the measurement a cell is located. The sharp increase of the errors just to the right of the measurement locations indicates that the Kalman gain only reflects a small width of update that is directly influenced by the measurement location. The increase in uncertainty moving to the right of the measurement location demonstrates how the EKF grows more uncertain the longer it propagates or the further from a measurement the solution is located. The L2 norm is equal to 0.719381, indicating that the filter is continuing to improve its solution with the increased process noise error. The final panel, Fig. 5.3(d), demonstrates a process noise covariance set equal to $Q = 10$. The uncertainty bounds are shown to be very large and the solution only improves minimally compared to that of $Q = 1$ with a L2 norm equal to 0.718288.

Studying the process noise covariance yields a similar change in solution results. Figure 5.4 demonstrates how the error bounds in the near region of the measured values grow with an increased uncertainty in the measurement. For this study, the process noise covariance is set as $Q = 0.1$ for all four cases. In Fig. 5.4(a), no measurement noise is considered. Thus, the uncertainty bounds are shown to nearly disappear at the measurement location. As some measurement noise is assumed, as in Fig. 5.4(b), slight increases in the uncertainty can be noted. Further in-

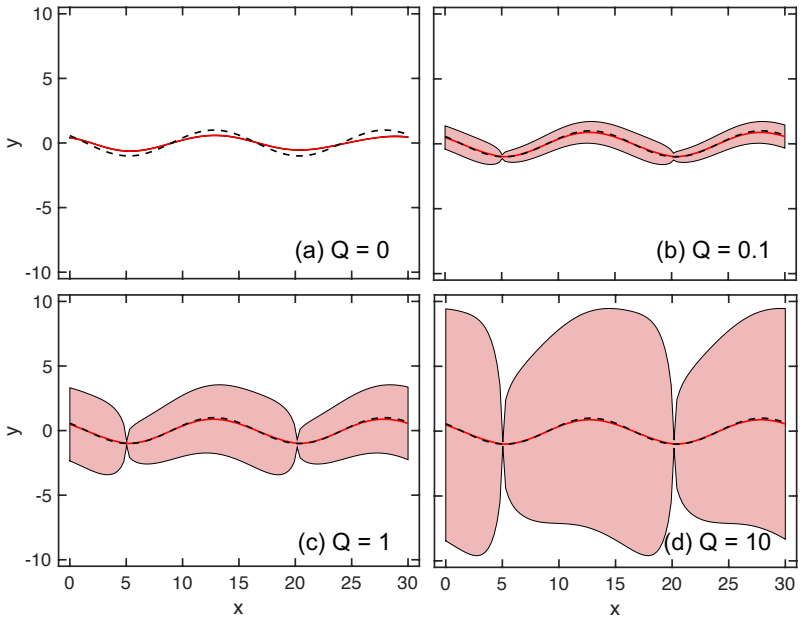

Figure 5.3: The EKF estimate (red solid line) using the reference dataset (black dashed line) at the final update for different process noise covariance values, Q, using the 3σ uncertainty bounds (pink shade). (a) $Q = 0$, (b) $Q = 0.01$, (c) $Q = 1$, and (d) $Q = 10$.

creasing the measurement error by an order of magnitude demonstrates a significant increase in uncertainty near the measurement locations. Because the measurement noise at $x = 20$ is considered twice as much as that at $x = 5$, the uncertainty bounds at $x = 20$ are noticeably larger in Fig. 5.4(c). With large enough measurement errors, the uncertainty at the location of the measurement can become nearly as large as the uncertainty throughout the remainder of the domain, as demonstrated in Fig. 5.4(d). Due to the increase in measurement uncertainty for a test case with no measurement error, the L2 norm is seen to increase from 0.718166 in Fig. 5.4(a) to 0.878925 in Fig. 5.4(d).

Changing the folding scale changes the initial growth and development of the solution and covariance bounds, but does not affect the final results. Further parameter sweeps were not considered as part of this study as the goal was simply to replicate the solutions of previous work [150] and gain a general understanding of how the estimates are affected by the noise covariances. Note that the cells that do not correspond to a measurement location in this model will never see a reduction in uncertainty. These results can be improved by using spectral diversity applying the

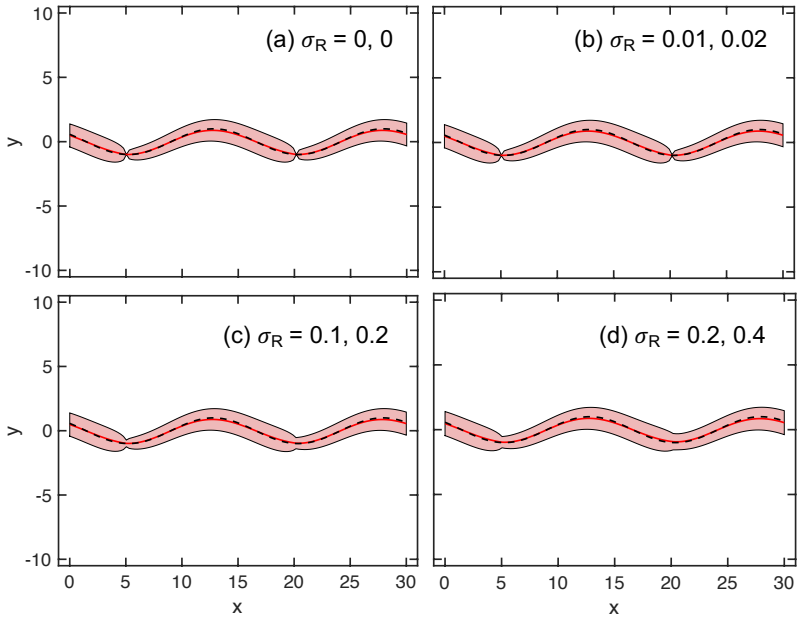

Figure 5.4: The EKF estimate (red solid line) using the reference dataset (black dashed line) at the final update for different measurement noise values, σ_R, using the 3σ uncertainty bounds (pink shade). The process noise Q is set as $Q = 1$.(a) $\sigma_R = 0, 0$. (b) $\sigma_R = 0.01, 0.02$ (c) $\sigma_R = 0.1, 0.2$ (d) $\sigma_R = 0.2, 0.4$

process noise in the modal space. This allows for a cell-to-cell mapping that benefits from constraints on the physics and only the considered parameters, rather than updating every element in the domain. This is reserved for future work as the application of state estimation techniques to one-dimensional systems is considered with greater detail.

5.2 One-Dimensional Thermal Conduction

In seeking a simple one-dimensional problem that better reflects a Hall effect thruster model, a thermal conduction test case is developed. Thermal conduction studies the transfer of internal energy by particle collisions and electron motion over some region of space. Heat flows from hotter areas to colder areas, indicated in this study by higher values dissipating as they are diffused towards the lower-value cells. The quantification of the ease with which a particular medium conducts heat is termed the thermal conductivity, κ. This material property is imperative to modeling how the system conducts heat, but is a relatively hidden parameter in the set of equations, much like electron mobility in Hall effect thruster models. Studies have been previously performed using

the EKF to estimate thermal conductivity parameters in other literature [151].

The general thermal conduction equation can be written as

$$u_t = \kappa u_{xx}, \tag{5.7}$$

where κ is the conductivity parameter and the subscripts t and x indicate partial derivatives with respect to time and space, respectively. The double xx subscript indicates a second derivative in space.

5.3 Simulation Setup

The initial condition for this thermal conduction case over 100 cells that discretize the domain from 0 to 1 is set as a discontinuous pulse in the center of the domain, occupying ten cells on either side of center such that

$$\begin{aligned} &\text{if}(x < 0.4 || x > 0.6) \rightarrow u = 0, \\ &\text{if}(0.39 < x < 0.61) \rightarrow u = 1. \end{aligned} \tag{5.8}$$

For this work, the system is propagated as first order forward in time and second order in space using a central differencing scheme as shown below

$$\frac{u_{k+1}^n - u_k^n}{\Delta t} = \kappa \frac{u_k^{n+1} - 2u_k^n + u_k^{n-1}}{\Delta x^2}. \tag{5.9}$$

This simple solution is entirely dependent on the form of the conductivity coefficient provided over time for each cell in the system as well as the boundary conditions. The thermal conductivity coefficient can be treated as a single constant, multi-region values, or as time-dependent solutions. For this work, both Neumann and Dirichlet boundary conditions are employed depending on the test case. These conditions either specify the slope of the boundary cells with the domain, or the value of the boundary cells, respectively.

5.3.1 EKF Setup

The EKF can be constructed similarly to the linear advection case. The initial condition given to the EKF is a zero-value domain with some guess for the initial thermal conductivity coefficient or coefficients. Again, an original EKF model is used with fixed values of Q and σ_R. The state vector consists of the solution at every cell as well as the conductivity coefficient.

Using the propagation equation given in Eq. (5.7) and the state vector of every cell value and thermal conductivity coefficient, the Jacobian equations are written as

$$\frac{d}{du_n} = \frac{-2\kappa}{\Delta x^2}, \tag{5.10a}$$

$$\frac{d}{du_{n+1}} = \frac{\kappa}{\Delta x^2}, \tag{5.10b}$$

$$\frac{d}{du_{n-1}} = \frac{\kappa}{\Delta x^2}, \tag{5.10c}$$

$$\frac{d}{d\kappa} = \frac{u_{n+1} - 2u_n + u_{n+1}}{\Delta x^2}. \tag{5.10d}$$

Neumann boundary conditions are set and thus present slightly modified versions of these equations at the first and last cells in the domain using only the existent adjacent cell to the left or right of the boundary cell, depending on which boundary is in question.

The covariance matrix is constructed the same way as for the linear advection test case using a set r_e folding scale to initialize the matrix. As the number of measurements increases, the Kalman gain uses the summation of the different measurement signals to generate the solution. The propagation is ensured to meet the Von Neumann condition as $\kappa\Delta t/\Delta x^2 \leq 1/2$ to ensure a stable solution. The EKF uses a time step of $\Delta t = 2.5 \times 10^{-5}$ with a measurement signal of $\Delta t_m = 2.5 \times 10^{-4}$.

5.3.2 Estimating Thermal Conductivity Coefficients

The simplest test case involves treating the domain from $x \in [0, 1]$ as a single region with one thermal conductivity coefficient. A parameter sweep was performed to compare the solutions using

only one or two measurement locations, as well as varying those measurement locations, as shown in Fig. 5.5. The solution is run for 0.35 units of time with a cell width of 0.01 and propagation time step of $\Delta t = 2.5 \times 10^{-5}$. The initialized domain is set to zero everywhere and $\kappa_0 = 0.9$. As with the linear advection case, constant values for the process noise and measurement noise covariances are set as $Q = 1$ and $\sigma_R = 0.01$. The folding scale is set at $r_e = 0.01$.

For the single measurement signal case, Fig. 5.5(a), the estimated domain is noted to favor and shift its peak towards that measured location. While only results using locations on the left hand side of the domain are presented, the trend was confirmed to continue as the measurement signals were placed further to the right hand side of the domain. This shift is the cause for uneven boundary conditions in nearly all cases presented, despite the uniform initial pulse and subsequent diffusion evident in the true solution. As expected, as the number of measurements increases, the overall estimate of the domain improves, as seen in Fig. 5.5(b). There is little discernible difference between the estimates for different measurement locations. Even when the measurement locations favor one side of the domain, such as $x = 0.45, 0.5$, the overall estimate of the domain is very close to the true solution.

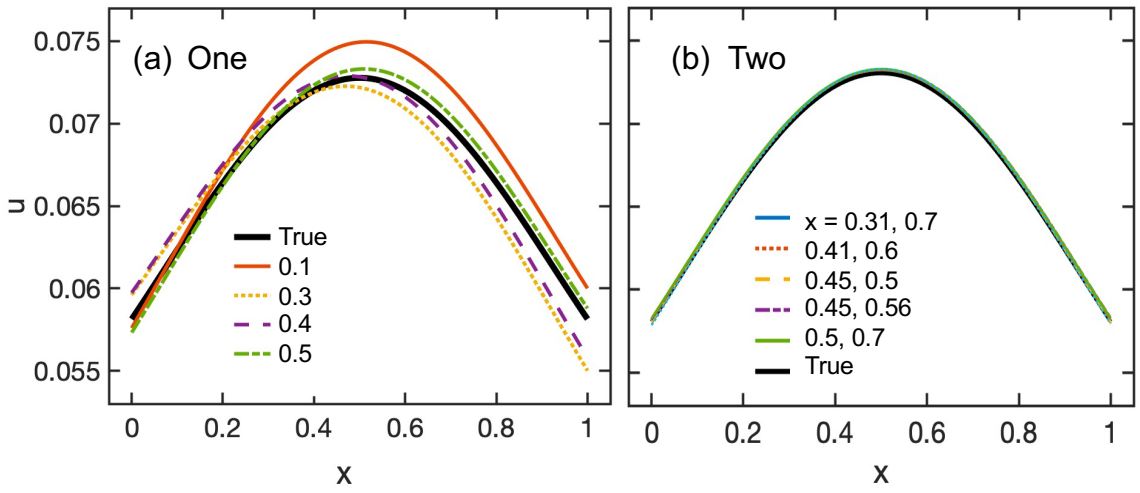

Figure 5.5: The EKF estimates for (a) one and (b) two measurement signals to estimate a single thermal conductivity coefficient, κ. The true solution is indicated by the black solid line.

Comparing the transient estimated values of the thermal conductivity coefficient reveals two key trends. The first is that the transient time where the filter searches for the correct estimate is significantly faster when using two measurement signals compared to one. The solutions in Fig. 5.6(a) can be seen to change over two-thirds of the total simulation time. In comparison, the estimates in Fig. 5.6(b) are seen to change very little after the first 0.002 time units. The second trend is that while the final estimates for the single measurement case show a very narrow standard deviation between different measurement locations, apart from the $x = 0.1$ measurement, the two measurement case has a discernibly wider spread of final estimates. Yet, the solutions in Fig. 5.5(b) when two measurement signals are used demonstrate better agreement with the reference diffusion solution in every case. The difference in estimate results can be attributed to the additional error incurred by starting the domain from a zero-valued condition whereas the reference solution starts from a discontinuous pulse. In this case, the second measurement is used by the EKF to compensate for the lack of knowledge regarding the initial condition of the domain as opposed to improving the estimate of the unknown thermal conductivity coefficient.

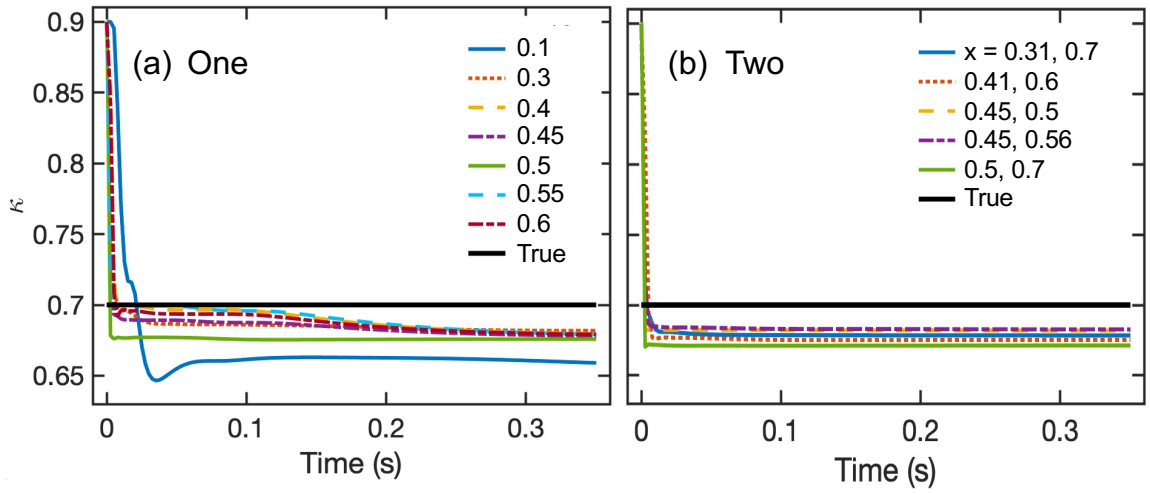

Figure 5.6: The EKF estimates for (a) one and (b) two measurement signals to estimate a single thermal conductivity coefficient, κ. The true solution is indicated by the black solid line. The numbers in the legend indicate the measurement locations

The L2 error norms for each measurement test case can also be plotted to compare the numerical accuracy of the final solutions. It is immediately evident from Fig. 5.7 that the estimates provided two measurement signals produce smaller errors than the estimates provided only one measurement signal. The two measurement cases are plotted based on the first measurement signal location, but due to the existence of two cases with a measurement at $x = 0.45$, one of these cases is plotted at the location $x = 0.56$. This study confirms the notion that increasing the number of measurement signals can improve the overall estimation of a system. Initial results using three measurement signals demonstrated only slight improvement in the thermal conductivity coefficient estimates.

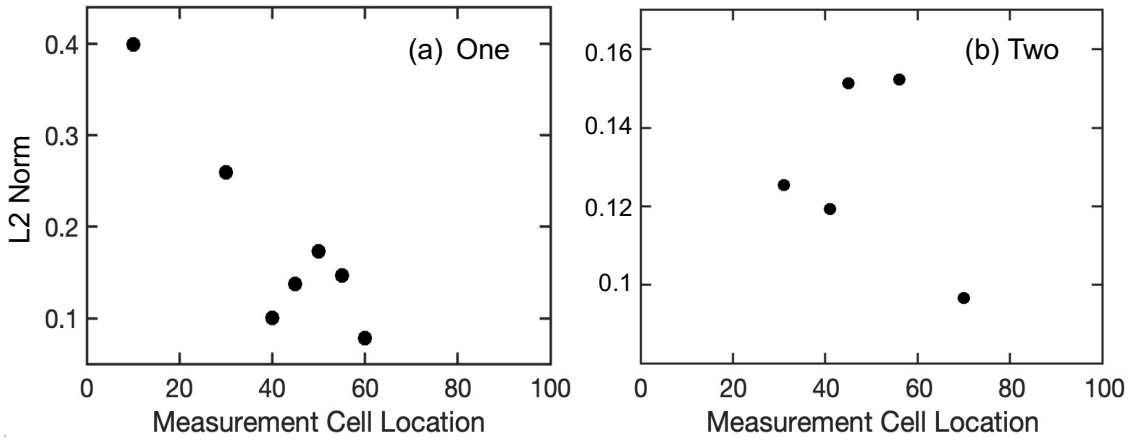

Figure 5.7: The L2 error norms for (a) one and (b) two measurement signals for estimating the domain and thermal conductivity coefficient. The two measurement cases are plotted based on the first measurement signal location, but due to the existence of two cases with the first measurement location at $x = 0.45$, one solution is plotted at the location $x = 0.56$.

The final study using a single thermal conductivity coefficient tracks the effect of refining the grid for one, two, and three measurement locations, shown in Fig. 5.8. The time history of the L2 error norm is plotted for these cases to demonstrate how the overall estimation process is affected by the refinement of the grid. Identical measurement locations were used for this test case. The finer grid required a smaller time step to meet the CFL condition, set as $\Delta t = 5 \times 10^{-6}$ for a

cell width of $\Delta x = 0.005$. All other operating conditions were kept the same between all plotted cases. It is noted that due to the doubling in size of the matrices, there is an exponential increase in computational cost for such this test case. While the 100 cell case runs in approximately 5 minutes, the 200 cell case takes nearly 2.5 hours. This is caused by the increased number of operations required for a matrix with 10,000 cells compared to a matrix with 40,000 elements. Furthermore, it is noted that the finer grid solutions do not reach as low of error values as the coarse grid. While the physics-based model can be better captured with a finer grid due to a reduction in error, the EKF solution does not share this trend. With more cells to match, the EKF is shown to perform worse for the finer grid in all cases because of the greater reliance on the cross-covariance terms to transfer information to other cells. This increase in the number of unknown cells with the same number of measurement locations requires an examination of the observability of the system to determine an appropriate number of measurements. Significantly, the single measurement case for the fine grid case did not produce feasible results, instead finding a wildly oscillating solution. For this reason, it is not plotted in this figure.

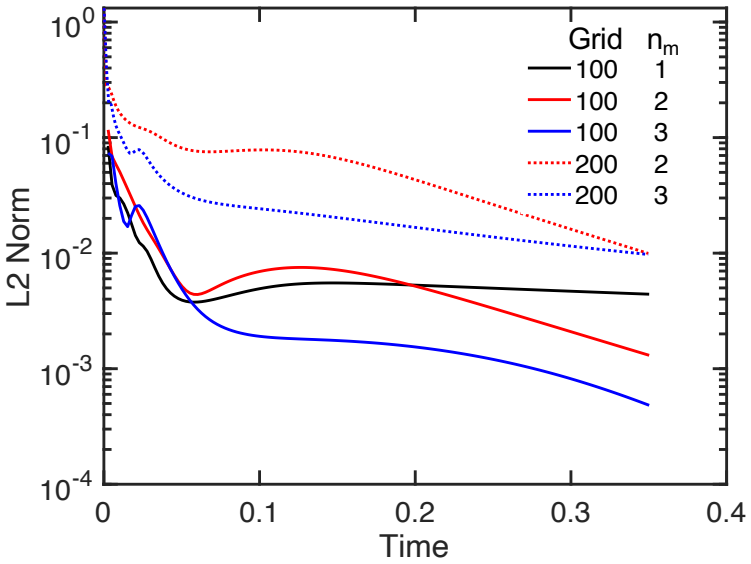

Figure 5.8: A comparison of the L2 error norms for different numbers of measurements. The solid lines indicate the results for 100 cells while the dotted lines indicate 200 cells. One (black), two (red), and three (blue) measurement locations are used by the EKF to estimate the true solution of the domain, corresponding to the n_m column of the legend.

Having thoroughly studied the applications of the EKF with a one-region case, a new reference solution is created that consists of two regions, each with a different thermal conductivity coefficient. The domain is bisected in the middle, u_{50}, for the following studies. While not shown as part of this study, the EKF was tested given a single measurement location and a single estimated conductivity coefficient at various points in the domain for a reference solution based on two thermal conductivity coefficients. Based on the location of the measurement signal, either in the high κ region or the low κ region, the EKF is able to recover estimates of the thermal conductivity coefficient in the corresponding region to the correct order of magnitude. While this is valuable insight to how the filter operates, this also indicates the necessity for a sufficient number of measurement signals to fully capture the unknowns of the system.

Estimating two thermal conductivity coefficients with two measurement signals is demonstrated in Fig. 5.9. Here, both the domain estimates, Fig. 5.9(a), and κ estimates, Fig. 5.9(b), are shown. Representative cases of a good and a poor estimate are presented to demonstrate how widely the results can vary. The most significant cause of the poor estimation for the $x = 0.3, 0.7$ case is due to the incorrect initial condition for the domain. If the same test is run with a correct initial pulse rather than a zero initial condition, the resulting estimate is significantly closer to the reference solution of $\kappa_1 = 0.7$ and $\kappa_2 = 0.07$. The solution with both measurement conditions inside of the initial pulse reveals a significant improvement in the overall estimation due to the immediate information regarding the initial condition of the domain available to the EKF. Though the larger κ value on the left hand side of the domain is overestimated, the smaller κ is estimated more accurately. Still, this test case yields a L2 error norm of 1.0099.

The final case considered for this test case was the inclusion of a source term for the conduction model. The model now reads

$$\frac{du}{dt} = \frac{d^2u}{dx^2} + Q_s, \tag{5.11}$$

where the last term on the right, Q_s, is a source term. The source term is included as a pulse for this study, such that a repeatable oscillation is generated in the model. A condition is set in the

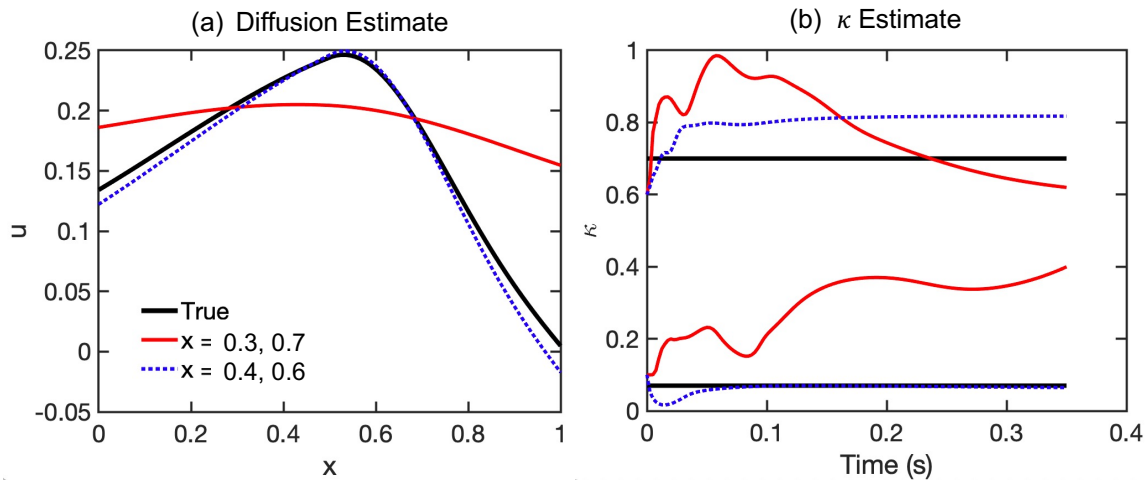

Figure 5.9: The estimation results for (a) the diffusing solution and (b) the κ estimates for two unknowns and two measurement signals. The $\kappa = 0.7$ value true solution corresponds to the left hand side of the domain, while the $\kappa = 0.07$ value is used for the right hand side of the diffusion solution. Note that both plots use the same legend.

reference heat diffusion model such that

$$\text{if } u_{50} < 0.6 \rightarrow Q_s = 0.001, \text{if } u_{50} > 0.8 \rightarrow Q_s = 0, \tag{5.12}$$

can be used to inform the value of the source term. Note that the value of the source term is set using these if-conditions until the alternate condition is met. Thus, once u_{50} is less than 0.6, until u_{50} is greater than 0.8, the source term is set as $Q_s = 0.001$. To achieve a repeatable oscillation, the boundary conditions were changed from Neumann to Dirichlet to keep the boundary values at a zero. Measurement signals are supplied for $x = 0.45, 0.55$ as it was discovered in previous studies that two measurement signals improved the overall accuracy of the estimate.

Figure 5.10(a) shows the value of the central cell in the domain to demonstrate effect of the source term, in red, compared to the sourceless diffusion case shown, in black. Note that the pulsed solution is perfectly repeated in each oscillation, but is only shown with one tenth of the original data in this figure, leading to an incomplete figure. Figure 5.10(b) compares the estimate of a single thermal conductivity coefficient, applied across the entire domain, for both the constant diffusion and the pulsed source term cases. The pulsed solution is seen to overestimate the true

thermal conductivity value of 0.7 as time continues, while the sourceless estimate has reached a near steady-state condition notably closer to the true solution. These results are the first indication that time-dependent oscillations affecting spatial distributions may be of concern in future applications of the extended Kalman filter. As the system has no knowledge of the pulse prior to the arrival of measurement data, the EKF attempts to account for the oscillations by altering the thermal conductivity coefficient. This incorrect assignment of the causation of the source term to a change in the thermal conductivity coefficient requires further study to determine whether certain limitations could be put on the EKF to improve the estimate.

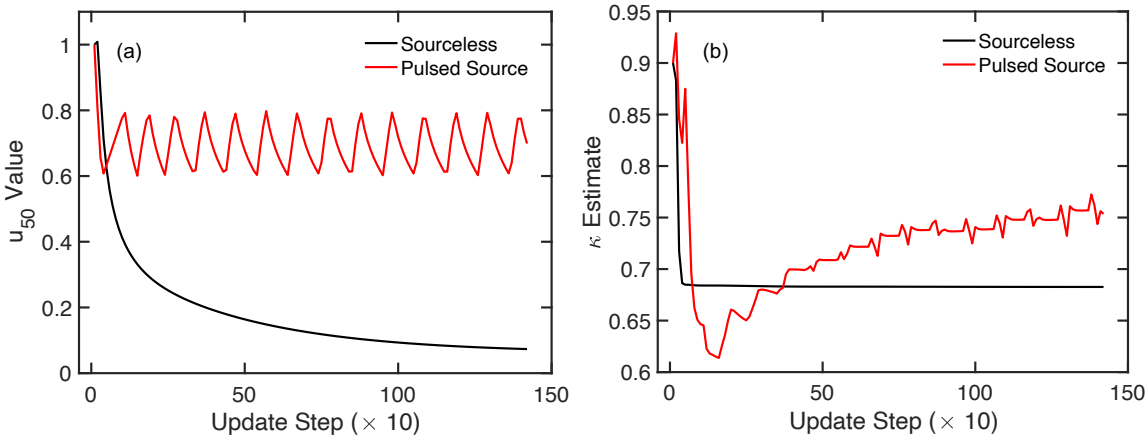

Figure 5.10: The estimation results for κ in an oscillating thermal conduction test case (red) in comparison to the sourceless case (black). (a) The value of the central cell in the domain, u_{50}. (b) The κ estimate.

The final goal of this study was to determine whether the boundary location between the two coefficients, i.e., where it changes, can be estimated. At a quick glance, there is no mathematical representation of this boundary location in the system of equations; it only exists as an if-statement in the physics-based model. For this reason, there are no Jacobian (covariance) terms that relate this boundary condition to the other states in the model. As the EKF is predicated on the cross-correlation of terms to inform the Kalman gain, the gain contains no update for the boundary term. Thus, with the current setup of the EKF, it is determined that estimation of the boundary between

solutions is not within the scope of this work.

5.4 Quasineutral Drift-Diffusion Model

Hall effect thruster research has improved tremendously over the last few decades, experimentally and computationally. Russia first experimented on HETs successfully in the 1960s, beginning with Morozov and leading to works by Smirnov and Zubkov, amongst many [152, 153]. A few decades later saw successful modeling work by Lentz and Fife that included effects such as quasineutrality, Bohm diffusion across magnetic field lines, fixed magnetic fields, and Maxwellian electrons [154]. Fife generated a model that tracked heavy particles with a PIC method while electrons were modeled as a fluid continuum [42]. Over the next two decades, leading to the present, experiments and models have seen vast improvement. New diagnostic tools such as laser-induced fluorescence [155], Laser Thomson scattering [51], and optical emission spectroscopy [156] have improved experimental measurements and enabled the study of smaller timescale phenomena. Each technique has its own advantages and disadvantages but collectively, these measurement techniques still fall short of reliably quantifying the high-speed, fast scale electron phenomena in the plasma discharge. Computationally, Fife's work has been extended and explored over the last two decades. Boeuf further studied the existence of breathing modes [28]. Koo and Boyd studied electron mobility coefficients and thruster erosion [32]. Hara used a hybrid direct-kinetic model to study mode transitions and a fluid model to study ionization oscillations [53, 157]. Others have used PIC models to analyze the length scales and frequencies needed for transport [41]. Throughout all of these studies, a few key challenges have been studied and identified.

Thruster erosion of channel walls has long been a leading factor in lifespan limitation [30]. This erosion is caused both by ion bombardment and possibly electron bombardment, which can be affected by the shape of the wall's decay over time [158, 159]. Secondary electron emission has been studied and revealed to influence the charge exchange [30]. The assumed electron bombardment has led to numerous studies of plasma oscillations and electron mobility. Recent work on magnetically-shielded HETs that use a refined magnetic field shape to help detatch the particles from the walls [160].

Electron mobility across magnetic field lines is still imperfectly modeled in state-of-the-art models [8]. This is often noted by a high electron current towards the thruster anode and higher electron densities noted experimentally compared to computational simulations [161]. Many forms of causation have been studied from electron-ion and electron-wall collisions to plasma oscillations near the exit [41, 65, 162]. As these dynamics are faster than reliable experimental data acquisition techniques, there is an added challenge to ensuring this phenomenon is correctly addressed in a model. Many state-of-the-art models rely on a Bohm mobility parameter and time-averaged quantities.

The third challenge is the simple difference between ground testing and flight performance, often termed facility effects [163]. A number of undesired effects have been noted to exist within the confines of a vacuum chamber, including back-sputtering of wall materials onto the thruster [164], electric charing of the walls [165], and changing background pressure in the chamber as the thruster runs [166]. Each of these affect thruster performance and give further need to develop new test methodologies or improve the computational modeling capabilities.

The Hall effect thruster used in this section is based on a one-dimensional quasineutral drift-diffusion (QDD) fluid model [167]. This model uses a quasineutral assumption for the plasma, a drift-diffusion approximation for the electrons, and accounts only for singly charged ions. Being based on a fluid approximation of Hall effect thruster operation, the conservation equations for mass, momentum, and energy can initially be written as

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\boldsymbol{u}) = S, \tag{5.13a}$$

$$\frac{\partial}{\partial t}(mn\boldsymbol{u}) + \nabla(mn\boldsymbol{u}.\boldsymbol{u} + p) = q(\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B}) + \nabla\tau + \boldsymbol{R}, \tag{5.13b}$$

$$\frac{\partial}{\partial t}(n\varepsilon) + \nabla \cdot (n\boldsymbol{u}\varepsilon + p\boldsymbol{u}) = \nabla \cdot \boldsymbol{Q} + qn\boldsymbol{u} \cdot \boldsymbol{E} + S_{elas} - S_{inelas} + \Phi, \tag{5.13c}$$

where n is the number density, $\boldsymbol{u}$ is the bulk velocity, ε is the mean energy, S is the source for particle density, p is the pressure tensor, $\boldsymbol{R}$ is the momentum transfer due to collisions, q is the heat flux vector, and S_{elas}, S_{inelas}, and Φ are energy exchange due to various types of collisions.

For the ion continuity equation, the S source term can be written as $S_{ion} = n_e \nu_{ion}$ to directly account for the effects of ionization where $\nu_{ion} = n_n k_{ion}$ is the ionization frequency based on the neutral number density n_n and ionization rate coefficient, k_{ion}. This rate coefficient is obtained from tabulated data generated by BOLSIG+ in the model [168].

For the drift-diffusion approximation, the conservation of momentum equation for electrons, assuming an axial direction, can be simplified by assuming that the transient and inertial terms are neglected since the model assumes electrons are at steady-state within the ion characteristic time and move at low Mach numbers. The resulting electron flux becomes

$$n_e \boldsymbol{u}_e = -n_e \bar{\bar{\mu}} \cdot \left(\boldsymbol{E} + \frac{1}{e n_e} \nabla p_e \right), \tag{5.14}$$

where n_e is the electron number density, $\bar{\bar{\mu}}$ is the electron mobility tensor, and p_e is the electron pressure. The electron pressure is considered isotropic and follows the ideal gas law. This leads to a one spatial dimension, two velocity dimension pair of equations to capture the electron transport as

$$\Gamma_{e,x} = n_e u_{e,x} = -\mu_{e,\perp} \left(n_e E_\perp + \frac{1}{e} \frac{\partial p_e}{\partial x}_\perp \right), \tag{5.15a}$$

$$\Gamma_{e,y} = n_e u_{e,y} = n_e u_{e,x} \Omega, \tag{5.15b}$$

where the cross-field electron mobility is considered to be $\mu_{e,\perp} = \mu_{\perp,clas} + \mu_{\perp,ano}$, the classical contribution to mobility is $\mu_{\perp,clas} = e/m_e \nu_m (1 + \Omega^2)^{-1}$, the anomalous contribution to mobility is $\mu_{\perp,ano} = \alpha/B$, the Hall parameter is defined as $\Omega = \omega_B / \nu_{m,e}$ based on the electron gyrofrequency $\omega_B = qB/m_e$, the momentum transfer collision frequency is ν_m, and α is an empirical coefficient for the anomalous electron transport. Note that the ideal gas law can be used to rewrite $p_e = n_e k_B T_e$ where k_B is the Boltzmann constant and T_e is the electron temperature in K.

The quasineutral assumption results in the use of a charge conservation equation to solve the electrostatic electric field. This equation can be written by taking the difference between the ion and electron conservation equations and assuming only electron-impact ionization from the ground

state neutral atom to singly charged ions to arrive at the form

$$\frac{\partial \sigma}{\partial t} + \nabla \cdot (e n_i \boldsymbol{u}_i - e n_e \boldsymbol{u}_e) = 0, \tag{5.16}$$

where $\sigma = e(n_i - n_e)$ is the charge density, n is the density of ions (subscript i) and electrons (subscript e), and $\boldsymbol{u}$ is the bulk velocity of ions and electrons dependent on the subscript. Note that the quasineutral assumption leads to $\sigma = 0$ and can be coupled with the charge conservation equation via the electron flux, Eq. (5.14), to derive an equation for the potential as

$$\frac{\partial \Gamma_{e,x}}{\partial x} = \frac{\partial}{\partial x}\left(\mu_{e,\perp} n_e \frac{\partial \phi}{\partial x} - \frac{\mu_{e,\perp}}{e}\frac{\partial p_e}{\partial x}\right) = \frac{\partial \Gamma_{i,x}}{\partial x}, \tag{5.17}$$

which provides a second-order partial differential equation for the electrostatic potential, ϕ, written in the cross-field direction. Note that the electric field is $\boldsymbol{E} = -\nabla\phi$ and the ideal gas law still applies. This equation can be solved using a tridiagonal matrix solver with Dirichlet and Neumann boundary conditions.

A further consequence of the quasineutral assumption is that the anode sheath is unable to be resolved by the model. Different studies have developed sheath boundary models to handle these simulations, and this work uses an ion-attracting, electron-repelling sheath that assumes a half-Maxwellian distribution of electrons. Using this information, the sheath potential at the anode can be calculated at every time step by

$$-\Gamma_{i,a} + \frac{1}{4} n_{e,a} \sqrt{\frac{8 k_B T_{e,a}}{\pi m_e}} \exp\left(-\frac{e V_a}{k_B T_{e,a}}\right) = \frac{j_d}{e}, \tag{5.18}$$

where j_d is the net current density and V_a is the anode sheath potential used as a boundary condition of Eq. (5.17). The second boundary condition assumed $\phi = 0$ at the quasineutral electron-injection plane at the cathode. The model interface quantities at the anode are denoted with a subscript a and are solved by extrapolating the values from the cell centers.

The elastic energy loss is negligible in HET discharge plasmas as the electron temperature is typically larger than 5 eV where inelastic collisions become the dominant electron energy loss

mechanisms. Thus, an alternative to Eq. (5.13c) for calculating the electron energy is to solve the internal energy equation. Subtracting the energy equation obtained from the conservation of momentum from the total energy equation, Eq. (5.13c), yields

$$\frac{\partial}{\partial t}\left(\frac{3}{2}n_e k_B T_e\right) + \nabla\cdot\left(\frac{5}{2}n_e k_B T_e \boldsymbol{u}_e + \boldsymbol{q}_e\right) = \nabla p_e \cdot \boldsymbol{u}_e + m_e n_e \nu_m |\boldsymbol{u}_e|^2 - S_{loss}, \tag{5.19}$$

which includes heating due to collisional drag. This equation is solved using a second-order implicit Crank-Nicolson scheme to integrate the left hand side in time while the right hand side is solved explicitly. For the model used in this dissertation, a tridiagonal solver is used to solve Eq. (5.19).

Pressure can be solved in a coupled fashion with momentum to reduce numerical oscillations in the model. Assuming quasineutrality, $n_e = n_i$, the electric field can be shown to be dependent on the ion density, leading to a nonlinear coupling between the electron pressure contribution of the density equation with the inviscid flux in the ion momentum equation. For this coupling, the ion momentum equation is written for non-magnetized, collisionless ions as

$$\frac{\partial(n_i \boldsymbol{u}_i)}{\partial t} + \nabla\left(n_i \boldsymbol{u}_i \boldsymbol{u}_i + \frac{p_i}{m_i}\right) = \frac{e}{m_i} n_i \boldsymbol{E}.$$

This ion momentum equation can be written by substituting

$$E = \frac{\boldsymbol{u}_e}{\mu_\perp} - \frac{1}{e n_i}\nabla(n_i k_B T_e),$$

to obtain

$$\frac{\partial(n_i u_i)}{\partial t} + \nabla\left(n_i u_i u_i + \frac{p_i + p_e}{m_i}\right) = -\frac{e n_i u_e}{m_i \mu_\perp}, \tag{5.20}$$

a modified ion momentum equation for the electron-pressure coupled method. Note that the right hand side is a function of $\boldsymbol{u}_e$ but recovers the electric field in the region where the effect of diffusion

flux is negligible. This direct coupling enables a smooth estimation of the electron pressure using standard inviscid flux calculations as described in the next paragraph. It is also noted that the equation agrees with the physics of the ion acoustic waves.

A Steger-Warming scheme is used to evaluate the left hand side of the ion conservation equations based on the inviscid Euler formulation. This scheme employs flux vector splitting by evaluating the positive and negative eigenvalues of the Jacobian matrix obtained from a general Euler equation. Cell-centered conservative quantities are reconstructed such that every cell interface consists of two values for each quantity. As this is not part of the work performed for this dissertation, further detail of the method can be found in Ref. 167.

5.4.1 Simulation Setup

The QDD model is solved on a coarse grid of 100 cells for a total time of 5 ms with a time step of $\Delta t = 4$ ns. Due to the nature of data collection in the original QDD model, a code-to-code verification measurement signal does not exist for the first 2.8 ms of the simulation. Thus, the model is run without the EKF for that length of time before the first measurement signal arrives. During this time, the electron mobility parameter is allowed to propagate as it would normally in the QDD model. At such a time when measurement data begins to arrive, the electron mobility parameter is held constant during QDD propagation phase and only updated during the measurement update phase. Measurement data includes a constant and sinusoidal discharge current trace created in MATLAB as well as the discharge current data taken from Ref. 3. For each measurement data, the measurement signal arrives every 5 μs.

The electron temperature at the anode and thruster exit is set as 3 eV. The discharge voltage is 300 V. The magnetic field is set as 151.2×10^{-4} Tesla. The thruster geometry is based on the well-known Stationary Plasma Thruster (SPT) 100 thruster with an inner radius of 0.0345 m, an outer radius of 0.05 m, the anode located at the left hand boundary of the domain, the channel exit at 0.025 m, and the domain exit at 0.05 m. The mass flowrate is assumed to be 5 mg/s while the inlet velocity is set at 270 m/s. The ion temperature is set to 0.095 eV, initially. The model assumes no secondary electron emission from the walls.

5.4.2 Estimating Electron Mobility with an Extended Kalman Filter

Initial efforts to incorporate the extended Kalman filter with the QDD model led to unphysical results. The original setup of the state vector and propagation scheme included the ion and neutral number density, electron temperature, and axial electron velocity as states to be estimated. The electron velocity was updated by the measurement data, and then was used to solve for the electron mobility using the relation

$$u_e = -\mu_\perp \left(E + \frac{1}{e_{ch} n_e} \frac{\partial p_e}{\partial x} \right), \tag{5.21}$$

and solving for the anomalous contribution to the electron mobility separate from the EKF update equations using the approximation

$$\mu_\perp \approx \frac{m \nu_{eff}}{eB^2}, \tag{5.22}$$

based on the assumption that the Hall parameter, Ω, is large, e.g., $\omega_c \gg \nu_{eff}$.

The original QDD model is solved for every time step as it is originally written and documented. The covariance matrix uses a simplified set of equations based on the predator-prey work performed in Ref. 169 to calculate the Jacobian matrix. This prevented the use of complex equations in the Jacobian matrix based on the true relation of the states in the QDD representation. To aid in the simplification of the EKF model, rather than attempting to handle covariance matrices for 400 states all at once, e.g., four states for each of the one hundred cells in the domain, each cell is estimated with an individual state vector and covariance matrix. The results are stored in large memory banks to be called when needed, resulting in one hundred individually-run EKFs within the simulation.

The discrepancies between the simplified global model equations used in the covariance matrix and the full quasineutral drift-diffusion model led to a plethora of numerical artifacts within the estimation setup. Switching between local and global trends requires particular consideration with weighting schemes and physical constraints beyond those of the EKF. Additional attempts to in-

clude more streamlined calculations of the electron mobility and different physical considerations in the model showed little improvement in the unphysical solutions being returned by the EKF.

5.4.3 Estimating Electron Mobility with a One-Equation Correction Scheme

The difficulties of applying the EKF as originally intended to a Hall effect thruster model prompted an attempt to verify that a simpler version of an inverse problem would be feasible for this application. Having the QDD model as well as numerous discharge current signal traces, the decision was made to create a true inverse problem between the discharge current and the electron mobility parameter. Using the relation

$$I_d = e_{ch}(n_i u_i - n_e u_e), \tag{5.23}$$

the electron velocity can be substituted by Eq. (5.21) and rewritten to calculate the electron mobility parameter. Using the assumption that the discharge current is the same for every cell in the domain, this inverse equation can then be solved for every individual cell of interest as

$$\mu_\perp = \frac{(\frac{I_d}{A_{inlet}} - e_{ch} n_i u_i)}{e_{ch} n_i (E + \frac{\partial p_e}{\partial x})}, \tag{5.24}$$

based on a measurement signal arriving every 0.5 μs. The electron mobility is solved on the cell-interfaces using Eq.(5.24) and a simple averaging scheme is used to translate the solved cell-interface values to the required cell-center values. The same averaging scheme is also used to calculate the interface values for the ion number density and ion velocity before solving for the electron mobility. No other states are updated in the QDD model, requiring the system to adjust to changes in electron mobility between two measurement data. Once the update scheme begins, the QDD model no longer propagates the electron mobility parameter, further forcing the physics-based model to adjust to the calculated mobility parameter.

The results of numerous studies indicate that numerical artifacts are seen to arise from the simple averaging scheme used in the model to transition between cell-center and cell-interface values. When supplying a constant measurement signal, the discharge current is shown to oscillate

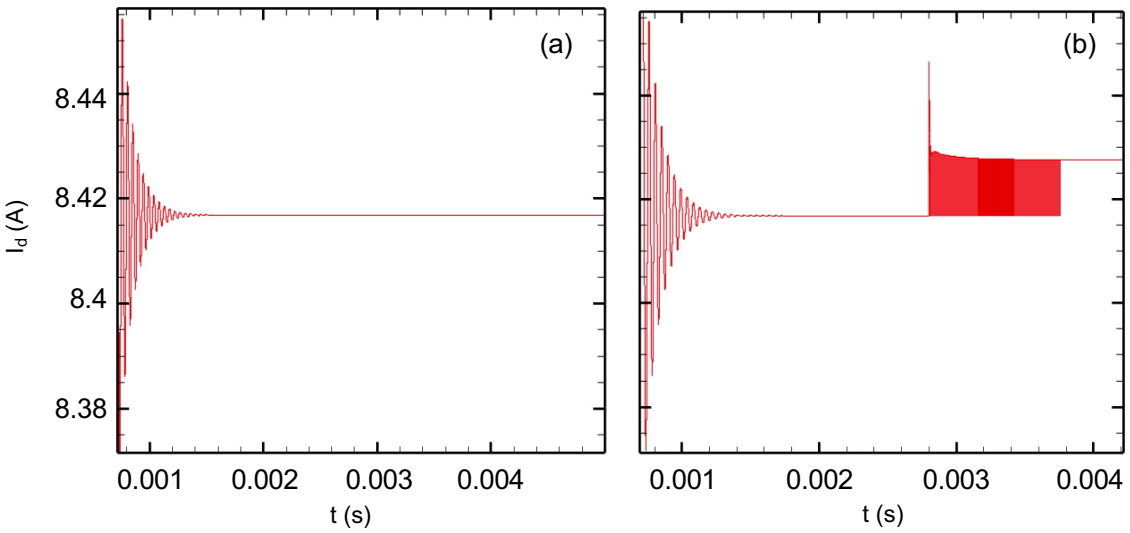

Figure 5.11: Example of how imperfections in double-point precision of the steady-state mobility solution changes the overall simulation results from (a) a steady-state condition to (b) a small amplitude oscillation.

around the desired solution due to the resulting non-constant mobility in time, as demonstrated in Fig. 5.11. Simple tests studying the transition from cell-center to cell-interface revealed numerical instabilities that caused oscillations, small in amplitude, to arise from the otherwise steady-state condition demonstrated in Fig. 5.11(a). After one averaging from cell-center to cell-interface and back, the reconstructed cell-center values do not match the original steady-state mobility to double-point precision and lead to a noticeable change in the discharge current, Fig. 5.11(b). The sensitivity of the QDD model to such minute changes in the electron mobility is the first indication of the litany of challenges that would be uncovered.

The fallout of this realization is that this particular setup for the electron mobility update is still too complex of a task. Using this knowledge, the scheme is changed to update specific regions of the domain while others are held constant, as shown in Fig. 5.12. Initial attempts of this selective update scheme follow the spirit of the three-region electron mobility model proposed in Ref. 33 that has become widely accepted. By selecting the region allowed to be updated, the results are intended to examine where the sensitive regions of the mobility profile exist. Starting by updating the channel, Fig. 5.12(b), the electron mobility values quickly begin to oscillate with large dis-

continuities that cause the entire simulation to fail due to the resulting negative discharge current values. These oscillations can be traced to the electric field, which is based on the gradient of the potential, ϕ. As the oscillations are noted to start in the near-anode region before growing, the understanding is that further considerations need to be made within the near-anode region of the domain. The two immediate factors to be considered are the application of an ion-attracting, electron-repelling sheath at the anode as well as the possible existence of nonmagnetized ions near the anode.

Moving to the exit of the channel, reveals a limited diffusion of the selected cells over the course of the simulation, Fig. 5.12(c). Much like the simple case of transferring values from cell-center to cell-interface and back, the resulting discharge current again demonstrates a small amplitude oscillation over the course of the correction scheme. When the plume is solved for, a discernible attempt to diffuse the value of the electron mobility parameter is visible in the time-dependent history of the parameter. This is noted in the difference of values after $x = 0.03$ m in Fig. 5.12(a) and Fig. 5.12(d) where the plume-updated values are notably lower than those of the steady-state mobility profile. Further testing revealed that this diffusion trend is followed when both the entire domain after $x = 0.023$ m is allowed to be updated. This diffusion does find a steady-state condition for the electron mobility value, but the discharge current remains oscillatory. As fewer cells are kept constant within the channel, the updates cells are seen to diffuse further. Although the discharge current trace is kept considerably constant in all of the cases presented in this section, there are still remaining questions about the lack of a true steady-state solution using the inverse problem.

The continually diffusing or oscillatory solutions recovered by this one-equation correction scheme indicate that further consideration is required to maintain better physical relevance for this estimation process.

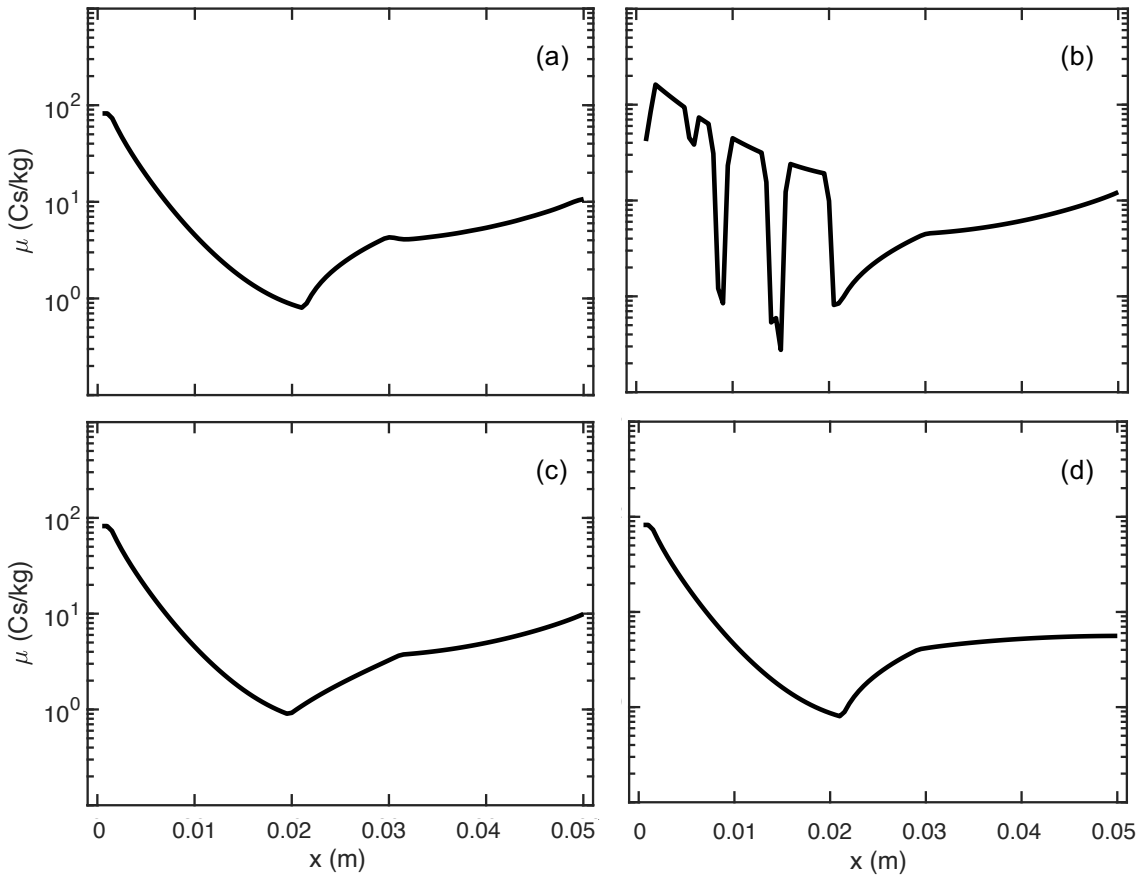

Figure 5.12: A selection of electron mobility solutions using the one-equation update. (a) The original, steady-state solution of the electron mobility. (b) The electron mobility if only channel cell values are updated, i.e., $x < 0.02$ m. (c) The electron mobility if only exit cells are updated, i.e., $0.02 < x < 0.03$ m. (d) The electron mobility is only the plume cells are updated, i.e., $x > 0.035$ m.

6. CONCLUSIONS

For my part I know nothing with any certainty, but the sight of the stars makes me dream.

- Vincent Van Gogh, *Arles, July 1888*

6.1 Summary

The complex interactions of plasma phenomena are known to greatly influence the behavior and performance of low-temperature plasma applications. Several numerical methods have been developed to study the detail of such physics to varying degrees of fidelity. One approach to combatting the increasing computational costs of high fidelity models is to use available experimental data as supplementary information to create a data-driven model. In this dissertation, a physically-constrained extended Kalman filter is developed for use with plasma physics models.

The development of the original EKF is presented to demonstrate the benefits of the physical constraints added in this work. Verification tests are performed using a chaotic Lorenz attractor and a driven-damped harmonic oscillator. The estimated results are shown to be in good numerical agreement with the reference solution physics. Studies with sparse measurements are used to demonstrate how the estimates change as less information is provided. The physically-constrained filter is shown to saturate at a certain order of magnitude of error as the measurements continue to become sparse. The filter is also shown, through an amplitude and phase study, to converge to the true dynamics with increasing measurement signal frequency for cases without measurement noise. Overall, these studies were also used to demonstrate the robustness of the physically-constrained EKF compared to the original EKF in the face of sparse measurement signals.

Further tests are run using experimental Hall effect thruster discharge current data with a simplified global xenon model. A two-species predator-prey model driven by the electron temperature fluctuations is used to demonstrate the physical relevance of the filter estimates. In particular, the dynamics of the incoming discharge current measurement signal are found in all states estimated

by the EKF. The filter is presented with two slightly different physics-based models to demonstrate the effect of the propagation model on the overall estimate. The inclusion of electronically excited states is shown to decrease the estimated electron temperature, as expected. The filter is demonstrated with different Hall effect thruster operation cases that range near steady-state operation to the highly oscillatory breathing mode operation. The physics-constrained EKF is demonstrated to robustly capture every case, unlike the original EKF that was unable to find a viable solution for the breathing mode oscillations.

The physics-constrained EKF is then applied to three plasma chemistry test cases of increasing complexity. Verification tests are performed using a global pure argon chemistry model to study the ability of the EKF to estimate more than one unknown at a time. The filter is shown to recognize the significance of different unknowns, giving preference to the more important states in the model. This preference is noticed in which states the EKF corrects over time, and which it does not alter. Cases are shown where unknown states of equal significance are both updated as well as a particular case where the physics-based propagation is so poorly defined that the system fails before the first measurement update. The filter shows the capability to estimate reaction rate coefficients, which are less obvious states in a system compared to the electron temperature. Having verified the model, experimental data from a pulsed inductively coupled plasma test are provided to the filter to determine the driving effective electron input power. A variety of peak powers, duty cycles, and pulse frequencies are studied. Results demonstrate that not only can the filter recognize a power on versus power off phase based on the provided ion number density measurement, but also the filter recognizes the difference in effective power required to generate the plasma versus sustain it. Physical relationships between the background ion density values and peak effective input power are shown to be consistent across test cases. It is further shown that only about one-third of the supplied power is estimated to be absorbed by the plasma in these models, likely due to the multitude of efficiency loss mechanisms at work within these plasma systems coupled with the use of a zero-dimensional model. Finally, an argon-oxygen plasma chemistry model is developed. Assuming that the electron power input is constant, the effective ion wall diffusion rates are

considered the unknown parameters to sufficiently capture the experimental data. Studies within this model indicate the benefits towards using a measurement signal for each element to capture a significant portion of the physics so that the EKF is correctly informed. Additionally, allowing the EKF more control over the state of the model is shown to improve the estimates of the system.

Having sufficiently tested the PC-EKF with global models, the EKF was expanded for use with models over one spatial dimension. An initial recreation of a linear advection propagation of a sinusoidal wave found in the literature is performed. This study is used to demonstrate that the initial domain of a system can be reconstructed using an EKF. An added benefit provided by the EKF is that so long as measurement data continues to be supplied to the EKF, the advection (dissipation) of the solution over time caused by using a first-order upwind propagation scheme is compensated for by the EKF. This exemplifies how the EKF can be used to improve the solutions of low-fidelity and low-dimensional systems.

A heat diffusion solution using a varying number of diffusion regions is also studied. This particular case is reflective of estimating an electron mobility coefficient in a Hall effect thruster that drives the motion of electrons within the system. The diffusion coefficient is an underlying parameter such as the mobility coefficient, and is as significant to the system. A variety of parametric studies changing the measurement locations, number of measurements, number of regions, grid refinement, and measurement frequency are performed to gain a better understanding of how the EKF behaves to improve future analysis with a true Hall effect thruster application.

Lastly, a series of initial studies using the extended Kalman filter and a one-equation inverse problem are applied to a fluid Hall effect thruster model using a quasineutral drift-diffusion formulation. The key assumptions made for this study include magnetized electrons near the anode, constant discharge current across the entire domain, and the ability of the QDD model to relax to a set electron mobility profile. This work demonstrates the sensitivity of the model to the profile of the electron mobility parameter, revealing effects of slight changes in the electron mobility values in the resulting discharge current. The select update ranges demonstrate that it is possible to use an inverse-problem type update without causing the simulation to fail if the channel profile is held

constant. Although the inverse problem returns diffusive solution for the electron mobility profile, adding a second updated quantity to the system or creating new physical constraints are promising options for improvement. Furthermore, as other studies stress the limitations of the extended Kalman filter, investigation of alternate state estimation techniques is reserved for future work.

6.2 Future Work

6.2.1 Improving the Physical Constraints

The present physical constraints are designed to ensure a positive-definite condition for the covariance of the unknown state. While these constraints remove the tight restrictions of the previous *ad hoc* parameters, they do still require some manual tuning, especially for the measurement noise covariance, R. Though this value can typically be set based on the known measurement error of the incoming data, there is more than one case where the assumed noise has to be set greater than the true noise to achieve close accuracy in the estimation. Also, because of the structure of the constraint, the condition is rarely met that requires an update to R. Thus, a new constraint on the measurement noise may be advisable to reduce the manual tuning required.

While the physics-based constraints developed as part of this dissertation focus on ensuring positive uncertainty bounds to prevent simulation failure, little work has been performed to develop physics-based constraints relevant to plasma physics. The purely numerical scheme of the extended Kalman filter leaves a significant gap for the inclusion of plasma physics knowledge that is more inherent to the state estimation technique.

6.2.2 Increasing Complexity

The physics-based models used throughout this dissertation are intentionally kept simple aside from the Hall effect thruster model. Particularly, the models have not estimated a sizable number of states to study the computational cost of the PC-EKF. The short simulation times of test cases such as the Lorenz system or a linear advection scheme require so little computational time that tracking the time spent on the EKF functions can be misleading. There is also a balance to be found between the complexity of the physics-based model compared to the extended Kalman filter. For

many of the cases in this dissertation, the EKF requires a larger number of equations to be solved than the physics-based model. Knowing this to be the case, time readouts were not included in any of the presented studies due to a disproportionate balance between the physics-based models and the filtering processes. Furthermore, some of the models in this dissertation were developed using MATLAB, which is a non-compiled language and operates much slower than languages such as C++. The application of the EKF to various systems can be investigated to determine if a balance exists between how many EKF equations are required and the physics-based model in a form of a scalability test.

6.2.3 Adding Detail to Zero Dimensional Hall Effect Thruster Model

The zero-dimensional Hall effect thruster model presented in this dissertation is based on a simplified predator-prey representation of ion and neutral number densities from Ref. 112. Though this model has been successfully utilized to describe the breathing mode oscillations observed in HETs, it fails to account for radial diffusion and neglects the dynamic behavior of electron temperature and velocity. A more representative set of equations would include an electron momentum equation that details various source terms for the physical processes and includes an electron velocity term. This additional equation enables the inclusion of an electron mobility parameter that can be estimated.

An alternative option is to shift the focus of the zero-dimensional model to include an electrical circuit. Even a simple resistor, inductor, and capacitor circuit can be included to develop a system of equations relating the thruster operation to the electrical circuit of the vacuum chamber. Such models are growing in popularity due to the realization that the behavior of thrusters in ground testing may not be the same as that observed during in-space operations. By adding detail to the Hall thruster physics-based model, past experimental work that has studied the coupling of HET discharge plasmas and the electrical configuration of the test facility can be introduced to the HET model through the EKF to study the effects on the physical processes [165, 60, 170].

6.2.4 Complex Plasma Chemistry Models

Having demonstrated that the EKF can be used to study reaction rate coefficients in a pure argon and mixed argon-oxygen inductively coupled plasma source, further research can be conducted with this knowledge. The gas mixtures can become more complex, including more reactions and states, while the number of unknowns or available measurement signals can be varied. Many reaction rates used in state-of-the-art models are either extrapolated from similar processes of other gasses or are set as constant values based on educated guesses. The EKF method can be applied in any number of these cases to verify the assumed reaction rate coefficients and improve confidence in these models.

6.2.5 Correlation Between Number of Regions and Number of Measurements

As the one-dimensional studies are presented, a general rule found is to have at least one measurement per region of interest to sufficiently capture the profile of the unknown estimates. Studying an increased grid refinement and more regions also proved to be increasingly challenging when the initial condition of the domain is set to a zero-value condition. It is likely that future Hall effect thruster studies would have a non-zero, yet possibly incorrect, initial condition. This may lead to better overall estimates despite still being an incorrect initial condition because it is less for the filter to correct with only a finite number of measurement signals.

6.2.6 Estimation of Multi-Region Boundary Location

One study that is mentioned in this dissertation focuses on estimating not only the values of the thermal conductivity coefficient in multiple regions, but also estimating the location of the boundary between regions. As noted in the text, this boundary exists as an if-statement in the physics-based models and therefore is not a visible quantity for the EKF to estimate. It is possible that an equation can be constructed to calculate this boundary within the physics-based model such that it can be estimated by the EKF.

6.2.7 Estimating Electron Mobility

The EKF propagation scheme used to track the covariances of the one-dimensional Hall effect thruster is a highly simplified global model approach to the physics being captured by the fluid model. This decision is made to simplify the error propagation as the global model propagation is expected to have larger errors than the fluid model, yet this creates a disconnect between the uncertainties and the states, which leads to imperfect measurement updates.

To better estimate the electron mobility, a higher-fidelity model should be developed for the Jacobian matrix used in the covariance matrix propagation scheme. This higher-fidelity model could still be a global model approach, but should account for more detailed physics in the system. Wall sheath interactions should be reintroduced to the energy balance equation. Secondary electron emission can also be included. If the desire is to continue using a global model approach, further interactions can be accounted for in the model to include reactions beyond ionization and excitation. Additional ions and excited states can also be included in the model.

6.2.8 Estimation of Other Phenomena

While this dissertation studies the use of the PC-EKF for electron mobility coefficients in a Hall thruster simulation, the filter can be used to study any number of physical phenomena. Instabilities and oscillations, fluctuating electric fields, facility background pressure effects, and modeling of the electrical circuit are only a handful of phenomena that can be studied using a Hall effect thruster model and the PC-EKF. Additionally, any assumed parameters within the simulation can be treated as unknown parameters to study whether their assumed values are valid.

These results indicate that further consideration must be taken into the setup of this problem. The idea to use an update in the mobility parameter may be insufficient to dictate the state of the system overall, as cited by some observability concerns related to extended Kalman filters [77]. A more stringent averaging technique may be required to reduce numerical artifacts in the model as values are transitioned from cell-center to cell-interface. Specific, physical limitations may be required in the near-anode region to mitigate the oscillations that arise based on considerations

such as the anode sheath potential [112]. Isolating the mobility further into its anomalous and classical components may also alleviate some of the numerical discrepancies caused by updating the entire mobility value.

6.2.9 Alternate State Estimation Techniques

The goal of this dissertation is to develop a real-time state estimation technique for use with plasma physics applications. In particular, an extended Kalman filter has been presented and tested with a variety of simplified plasma physics models to demonstrate its capabilities as an estimation method that simultaneously captures the uncertainties of the system. Widely noted in literature are the limitations, computational cost, and challenges of applying the EKF to complex and highly nonlinear systems. While the creation of a pair of physics-based constraints has helped mitigate some of the challenges of applying the EKF to sensitive systems, there are alternative options to the EKF.

Within the Kalman filtering family exist both a particle filter known as the ensemble Kalman filter (EnKF) as well as a deterministic sampling approach called the unscented Kalman filter (UKF). Both filters remove the need to capture a covariance matrix to study the uncertainties of a system, eliminating all associated challenges with that representation. By using either a population of random points or carefully selected sample points, the values can be propagated through the true nonlinear system and capture higher order accuracy (compared to the EKF accuracy which is only first-order accurate). Though the EnKF is known to be computationally expensive, the UKF typically operates under a similar computational complexity to the EKF.

Outside of the Kalman filter, estimation techniques such as Gaussian collocation, Monte-Carlo based particle filters, and even physics-constrained neural networks. In a general collocation method, the state and control are discretized at a set of appropriately chosen points in the time interval of interest. The continuous-time optimal control problem is transcribed to a finite-dimensional nonlinear programming problem which is solved using available software. In a Gaussian quadrature orthogonal collocation method, the state is approximated using a basis of either Lagrange or Chebyshev polynomials, and the dynamics are collocated at points associated with a Gaussian

quadrature. Monte Carlo particle filters represent the posterior distribution of the state variables by a system of particles, instead of a state vector, which evolves as new measurements arrive. In practice, large numbers of particles may be required to provide viable solutions. Physics-constrained neural networks, or deep learning, methods incorporate the governing equations of the physical model in the loss/likelihood functions of the learning framework to generate solutions or relationships between inputs and outputs that remain physically relevant. All of these methods, as well as numerous others, have been documented in literature, though very few for plasma physics phenomena.

REFERENCES

[1] G. A. Hebner and C. B. Fleddermann, "Characterization of pulse-modulated inductively coupled plasmas in argon and chlorine," *Journal of Applied Physics*, vol. 82, no. 6, pp. 2814–2821, 1997.

[2] Y. Wang and J. Olthoff, "Ion energy distributions in inductively coupled radio-frequency discharges in argon, nitrogen, oxygen, chlorine, and their mixtures," *Journal of Applied Physics*, vol. 85, pp. 6358–6365, 1999.

[3] N. Gascon, M. Dudeck, and S. Barral, "Wall material effects in stationary plasma thrusters. i. parametric studies of an spt-100," *Physics of Plasmas*, vol. 10, no. 10, pp. 4123–4136, 2003.

[4] D. Goebel and I. Katz, *Fundamentals of Electric Propulsion: Ion and Hall Thrusters*. John Wiley & Sons, 2008.

[5] A. Brun, S. Turck-Chieze, and P. Morel, "Standard solar models in the light of new helioseismic constraints. i. the solar core," *The Astrophysical Journal*, vol. 506, no. 2, p. 913, 1998.

[6] G. Matthews, "Tokamak plasma diagnosis by electrical probes," *Plasma physics and controlled fusion*, vol. 36, no. 10, p. 1595, 1994.

[7] D. Beauchemin, "Inductively coupled plasma mass spectrometry," *Analytical chemistry*, vol. 80, no. 12, pp. 4455–4486, 2008.

[8] K. Hara, "An overview of discharge plasma modeling for hall effect thrusters," *Plasma Sources Science and Technology*, vol. 28, no. 4, p. 044001, 2019.

[9] J. Brophy, C. Garner, B. Nakazono, M. Marcucci, M. Henry, and D. Noon, "The ion propulsion system for dawn," in *39th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, pp. AIAA–2003–4542, 2003.

[10] P. Saikia, H. Bhuyan, M. Escalona, M. Favre, R. S. Rawat, and E. Wyndham, "A nonlinear global model of single frequency capacitively coupled plasma and its experimental valida-

tion," *AIP Advances*, vol. 8, no. 4, p. 045113, 2018.

[11] A. A. Fridman and G. G. Friedman, *Plasma medicine*. John Wiley & Sons Chichester, UK, 2013.

[12] D. B. Graves, "Plasma processing in microelectronics manufacturing," *AIChE Journal*, vol. 35, no. 1, pp. 1–29, 1989.

[13] R. Wolf and A. C. Sparavigna, "Role of plasma surface treatments on wetting and adhesion," *Engineering*, vol. 2, no. 06, p. 397, 2010.

[14] M. Konuma, *Film deposition by plasma techniques*, vol. 10. Springer Science & Business Media, 2012.

[15] S. Ali, I. A. Shah, A. Ahmad, J. Nawab, and H. Huang, "Ar/o2 plasma treatment of carbon nanotube membranes for enhanced removal of zinc from water and wastewater: a dynamic sorption-filtration process," *Science of the total environment*, vol. 655, pp. 1270–1278, 2019.

[16] N. Nakamura, H. Baba, Y. Sakaguchi, and Y. Kato, "Boron removal in molten silicon by a steam-added plasma melting method," *Materials Transactions*, vol. 45, no. 3, pp. 858–864, 2004.

[17] P. Boumans, R. McKenna, and M. Bosveld, "Analysis of the limiting noise and identification of some factors that dictate the detection limits in a low-power inductively coupled argon plasma system," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 36, no. 11, pp. 1031–1058, 1981.

[18] P. L. G. Ventzek, R. J. Hoekstra, and M. J. Kushner, "Two-dimensional modeling of high plasma density inductively coupled sources for materials processing," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 12, no. 1, pp. 461–477, 1994.

[19] D. Pröfrock and A. Prange, "Inductively coupled plasma-mass spectrometry (icp-ms) for quantitative analysis in environmental and life sciences: a review of challenges, solutions, and trends," *Applied spectroscopy*, vol. 66, no. 8, pp. 843–868, 2012.

[20] T. Lafleur, P. Chabert, and J.-P. Booth, "Electron heating in capacitively coupled plasmas

revisited," *Plasma Sources Science and Technology*, vol. 23, no. 3, p. 035010, 2014.

[21] K. Racka-Szmidt, B. Stonia, J. Zelazko, M. Filipiak, and M. Sochacki, "A review: Inductively coupled plasma reactive ion etching of silicon carbide," *Materials (Basel)*, vol. 15, no. 1, 2022.

[22] H.-C. Lee, "Review of inductively coupled plasmas: Nano-applications and bistable hysteresis physics," *Applied Physics Reviews*, vol. 5, no. 1, p. 011108, 2018.

[23] G. Chen and L. L. Raja, "Fluid modeling of electron heating in low-pressure, high-frequency capacitively coupled plasma discharges," *Journal of Applied Physics*, vol. 96, no. 11, pp. 6073–6081, 2004.

[24] F. Kr"uger, H. Lee, S. K. Nam, and M. J. Kushner, "Electric field reversals resulting from voltage waveform tailoring in ar/o_2 capacitively coupled plasmas sustained in asymmetric systems," *Plasma Sources Science and Technology*, vol. 30, p. 085002, aug 2021.

[25] C. Charles, "Plasmas for spacecraft propulsion," *Journal of Physics D: Applied Physics*, vol. 42, no. 16, p. 163001, 2009.

[26] S. Mazouffre, "Electric propulsion for satellites and spacecraft: established technologies and novel approaches," *Plasma Sources Science and Technology*, vol. 25, no. 3, p. 033002, 2016.

[27] S. Barral, J. Miedzik, and E. Ahedo, "A model for the active control of low frequency oscillations in hall thrusters," in *44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, pp. AIAA–2008–4632, American Institute of Aeronautics and Astronautics, 2008.

[28] J. P. Boeuf and L. Garrigues, "Low frequency oscillations in a stationary plasma thruster," *Journal of Applied Physics*, vol. 84, no. 7, pp. 3541–3554, 1998.

[29] E. Y. Choueiri, "Plasma oscillations in hall thrusters," *Physics of Plasmas*, vol. 8, pp. 1411–1426, 2001.

[30] M. D. Campanell, A. V. Khrabrov, and I. D. Kaganovich, "General cause of sheath instability identified for low collisionality plasmas in devices with secondary electron emission," *Phys. Rev. Lett.*, vol. 108, p. 235001, 2012.

[31] D. Sydorenko, A. Smolyakov, I. Kaganovich, and Y. Raitses, "Modification of electron velocity distribution in bounded plasmas by secondary electron emission," *IEEE Transactions on Plasma Science*, vol. 34, pp. 815–824, June 2006.

[32] J. W. Koo and I. D. Boyd, "Modeling of anomalous electron mobility in hall thrusters," *Physics of Plasmas*, vol. 13, no. 3, p. 033501, 2006.

[33] R. R. Hofer, I. Katz, I. G. Mikellides, D. M. Goebel, K. K. Jameson, R. M. Sullivan, and L. K. Johnson, "Efficacy of electron mobility models in hybrid-pic hall thruster simulations," in *44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, (Hartford, CT), pp. AIAA–2008–4924, July 2008.

[34] I. G. Mikellides, A. L. Ortega, I. Katz, and B. A. Jorns, "Hall2de simulations with a first-principles electron transport model based on the electron cyclotron drift instability," in *52nd AIAA/SAE/ASEE Joint Propulsion Conference*, (Salt Lake City, UT), pp. AIAA–2016–4618, 2016.

[35] R. Sahu, A. R. Mansour, and K. Hara, "Full fluid moment model for low temperature magnetized plasmas," *Physics of Plasmas*, vol. 27, no. 11, p. 113505, 2020.

[36] I. G. Mikellides, I. Katz, R. R. Hofer, and D. M. Goebel, "Magnetic shielding of a laboratory hall thruster. i. theory and validation," *Journal of Applied Physics*, vol. 115, no. 4, p. 043303, 2014.

[37] D. M. Goebel, R. R. Hofer, I. G. Mikellides, I. Katz, J. E. Polk, and B. N. Dotson, "Conducting wall hall thrusters," *IEEE Transactions on Plasma Science*, vol. 43, no. 1, pp. 118–126, 2015.

[38] P. N. Wainman, M. A. Lieberman, A. J. Lichtenberg, R. A. Stewart, and C. Lee, "Characterization at different aspect ratios (radius/length) of a radio frequency inductively coupled plasma source," *Journal of Vacuum Science & Technology A*, vol. 13, no. 5, pp. 2464–2469, 1995.

[39] M.-H. Lee and C.-W. Chung, "Self-consistent global model with multi-step ionizations in inductively coupled plasmas," *Physics of Plasmas*, vol. 12, pp. 073501–073501, 06 2005.

[40] J. T. Gudmundsson, "On the effect of the electron energy distribution on the plasma parameters of an argon discharge: a global (volume-averaged) model study," *Plasma Sources Science and Technology*, vol. 10, no. 1, pp. 76–81, 2001.

[41] T. Lafleur, S. D. Baalrud, and P. Chabert, "Theory for the anomalous electron transport in hall effect thrusters. ii. kinetic model," *Physics of Plasmas*, vol. 23, no. 5, p. 053503, 2016.

[42] J. M. Fife, *Hybrid-PIC Modeling and Electrostatic Probe Survey of Hall Thrusters*. PhD thesis, MIT, 1998.

[43] I. G. Mikellides and I. Katz, "Numerical simulations of hall-effect plasma accelerators on a magnetic-field-aligned mesh," *Phys. Rev. E*, vol. 86, p. 046703, Oct 2012.

[44] J.-P. Boeuf and L. Garrigues, "E×b electron drift instability in hall thrusters: Particle-in-cell simulations vs. theory," *Physics of Plasmas*, vol. 121, no. 1, p. 011101, 2018.

[45] C. M. Lam, E. Fernandez, and M. A. Cappelli, "A 2-d hybrid hall thruster simulation that resolves the$\boldsymbol{E} \times \boldsymbol{B}$ electron drift direction," *IEEE Transactions on Plasma Science*, vol. 43, no. 1, pp. 86–94, 2015.

[46] F. Taccogna and P. Minelli, "Three-dimensional particle-in-cell model of hall thruster: The discharge channel," *Physics of Plasmas*, vol. 25, no. 6, p. 061208, 2018.

[47] F. Cichocki, A. Domínguez-Vázquez, M. Merino, and E. Ahedo, "Hybrid 3d model for the interaction of plasma thruster plumes with nearby objects," *Plasma Sources Science and Technology*, vol. 26, p. 125008, nov 2017.

[48] T. Tajima, *Computational Plasma Physics*. Westview Press, 2004.

[49] X.-D. Liu, S. Osher, and T. Chan, "Weighted essentially non-oscillatory schemes," *Journal of Computational Physics*, vol. 115, no. 1, pp. 200–212, 1994.

[50] K. Hara, *Development of Grid-Based Direct Kinetic Method and Hybrid Kinetic-Continuum Modeling of Hall Thruster Discharge Plasma*. PhD thesis, University of Michigan, 2015.

[51] K. Hara and S. Tsikata, "Cross-field electron diffusion due to the coupling of drift-driven microinstabilities," *Phys. Rev. E*, vol. 102, p. 023202, Aug 2020.

[52] I. Katz and I. G. Mikellides, "Neutral gas free molecular flow algorithm including ionization

and walls for use in plasma simulations," *Journal of Computational Physics*, vol. 230, no. 4, pp. 1454–1464, 2011.

[53] K. Hara, M. J. Sekerak, I. D. Boyd, and A. D. Gallimore, "Mode transition of a hall thruster discharge plasma," *Journal of Applied Physics*, vol. 115, no. 20, p. 203304, 2014.

[54] K. Hara and K. M. Hanquist, "Test cases for grid-based direct kinetic modeling of plasma flows," *Plasma Sciences Sources and Technology*, vol. 27, p. 065004, 2018.

[55] J. Bareilles, G. J. M. Hagelaar, L. Garrigues, C. Boniface, J. P. Boeuf, and N. Gascon, "Critical assessment of a two-dimensional hybrid hall thruster model: Comparisons with experiments," *Physics of Plasmas*, vol. 11, pp. 3035–3046, June 2004.

[56] C. M. Lam, E. Fernandez, and M. A. Cappelli, "A 2-d hybrid hall thruster simulation that resolves the exb electron drift direction," *IEEE Transactions on Plasma Science*, vol. 43, no. 1, pp. 86–94, 2015.

[57] K. Hara and I. D. Boyd, "Axial-azimuthal hybrid-direct kinetic simulation of hall effect thrusters," in *34th International Electric Propulsion Conference*, (Kobe, Japan), IEPC-2015-286, July 2015.

[58] S. Yoshikawa and D. J. Rose, "Anomalous diffusion of a plasma across a magnetic field," *The Physics of Fluids*, vol. 5, no. 3, pp. 334–340, 1962.

[59] G. S. Janes and R. S. Lowder, "Anomalous electron diffusion and ion acceleration in a low?density plasma," *The Physics of Fluids*, vol. 9, no. 6, pp. 1115–1123, 1966.

[60] J. A. Walker, J. D. Frieman, M. L. R. Walker, V. Khayms, D. King, and P. Y. Peterson, "Electrical facility effects on hall-effect-thruster cathode coupling: Discharge oscillations and facility coupling," *Journal of Propulsion and Power*, vol. 32, no. 4, pp. 844–855, 2016.

[61] M. R. Nakles and W. A. Hargus, "Background pressure effects on ion velocity distribution within a medium-power hall thruster," *Journal of Propulsion and Power*, vol. 27, no. 4, pp. 737–743, 2011.

[62] C. D. Himmel and G. S. May, "Advantages of plasma etch modeling using neural networks over statistical techniques," *IEEE Transactions on Semiconductor Manufacturing*, vol. 6,

no. 2, pp. 103–111, 1993.

[63] F. Krúger, T. Gergs, and J. Trieschmann, "Machine learning plasma-surface interface for coupling sputtering and gas-phase transport simulations," *Plasma Sources Science and Technology*, vol. 28, no. 3, p. 035002, 2019.

[64] S. Guessasma, G. Montavon, and C. Coddet, "Modeling of the aps plasma spray process using artificial neural networks: basis, requirements and an example," *Computational Materials Science*, vol. 29, no. 3, pp. 315–333, 2004.

[65] B. Jorns, "Predictive, data-driven model for the anomalous electron collision frequency in a hall effect thruster," *Plasma Sources Science and Technology*, vol. 27, no. 10, p. 104007, 2018.

[66] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *Journal of Computational Dynamics*, vol. 1, p. 391, 2014.

[67] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations," *arXiv:1711.10561*, 2017.

[68] C. M. Greve, K. Hara, R. S. Martin, D. Q. Eckhardt, and J. W. Koo, "A data-driven approach to model calibration for nonlinear dynamical systems," *Journal of Applied Physics*, vol. 125, no. 24, p. 244901, 2019.

[69] A. Piccione, J. Berkery, S. Sabbagh, and Y. Andreopoulos, "Physics-guided machine learning approaches to predict the ideal stability properties of fusion plasmas," *Nuclear Fusion*, vol. 60, p. 046033, mar 2020.

[70] B. Adams, W. Bohnhoff, K. Dalbey, M. Ebeida, J. Eddy, M. Eldred, J. Frye, G. Geraci, R. Hooper, P. Hough, K. Hu, J. Jakeman, M. Khalil, K. Maupin, J. Monschke, E. Ridgway, A. Rushdi, J. Stephens, L. Swiler, D. Vigil, T. Wildey, and J. Winokur, "Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 6.8 user's manual," tech. rep., 2014.

[71] S. S. LLC, "Galaxy simulation builder users manual v6.7," tech. rep., 2018.

[72] L. A. Mcgee and S. F. Schmidt, "Discovery of the kalman filter as a practical tool for aerospace and industry." Technical memorandum for NASA, 1985.

[73] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME Journal of Basic Engineering*, vol. 82 (Series D), pp. 35–45, 1960.

[74] S. F. SCHMIDT, "Application of state-space methods to navigation problems," vol. 3 of *Advances in Control Systems*, pp. 293–340, Elsevier, 1966.

[75] M. Joerger and B. Pervan, "Kalman filter-based integrity monitoring against sensor faults," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 2, pp. 349–361, 2013.

[76] C. D. Petersen, R. Fraanje, B. S. Cazzolato, A. C. Zander, and C. H. Hansen, "A kalman filter approach to virtual sensing for active noise control," *Mechanical Systems and Signal Processing*, vol. 22, no. 2, pp. 490–508, 2008.

[77] J. Crassidis and J. Junkins, *Optimal Estimation of Dynamic Systems*. Chapman and Hall, 2004.

[78] M. W. Czabaj, M. L. Riccio, and W. W. Whitacre, "Numerical reconstruction of graphite/epoxy composite microstructure based on sub-micron resolution x-ray computed tomography," *Composites Science and Technology*, vol. 105, pp. 174–182, 2014.

[79] S. Schroeder, P. Preusse, M. Ern, and M. Riese, "Gravity waves resolved in ecmwf and measured by saber," *Geophysical research letters*, vol. 36, no. 10, 2009.

[80] S. Chen, "Kalman filter for robot vision: a survey," *IEEE Transactions on industrial electronics*, vol. 59, no. 11, pp. 4409–4420, 2011.

[81] T. Kobayashi and D. L. Simon, "Application of a bank of kalman filters for aircraft engine fault diagnostics," in *Turbo Expo: Power for Land, Sea, and Air*, vol. 36843, pp. 461–470, 2003.

[82] Y. Liu, X. Fan, C. Lv, J. Wu, L. Li, and D. Ding, "An innovative information fusion method with adaptive kalman filter for integrated ins/gps navigation of autonomous vehicles," *Mechanical Systems and Signal Processing*, vol. 100, pp. 605–616, 2018.

[83] B. D. Tapley, B. E. Schutz, and G. H. Born, *Statistical orbit determination.* Elsevier Academic Press, 2004.

[84] S. Bolognani, L. Tubiana, and M. Zigliotto, "Extended kalman filter tuning in sensorless pmsm drives," *IEEE Transactions on Industry Applications*, vol. 39, no. 6, pp. 1741–1747, 2003.

[85] V. A. Bavdekar, A. P. Deshpande, and S. C. Patwardhan, "Identification of process and measurement noise covariance for state and parameter estimation using extended kalman filter," *Journal of Process Control*, vol. 21, no. 4, pp. 585–601, 2011.

[86] R. Mehra, "Approaches to adaptive filtering," *IEEE Transactions on Automatic Control*, vol. 17, no. 5, pp. 693–698, 1972.

[87] S. Akhlaghi, N. Zhou, and Z. Huang, "Adaptive adjustment of noise covariance in kalman filter for dynamic state estimation," in *2017 IEEE Power Energy Society General Meeting*, pp. 1–5, 2017.

[88] S. Ungarala, E. Dolence, and K. Li, "Constrained extended kalman filter for nonlinear state estimation," *IFAC Proceedings Volumes*, vol. 40, no. 5, pp. 63–68, 2007. 8th IFAC Symposium on Dynamics and Control of Process Systems.

[89] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303–1318, 2010.

[90] N. Gupta and R. Hauser, "Kalman filtering with equality and inequality state constraints," *arXiv:0709.2791*, 2007.

[91] R. Mehra, "On the identification of variances and adaptive kalman filtering," *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175–184, 1970.

[92] C. O. Weiss and J. Brock, "Evidence for lorenz-type chaos in a laser," *Phys. Rev. Lett.*, vol. 57, pp. 2804–2806, Dec 1986.

[93] N. Hemati, "Strange attractors in brushless dc motors," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 41, no. 1, pp. 40–45, 1994.

[94] D. Poland, "Cooperative catalysis and chemical chaos: a chemical model for the lorenz

equations," *Physica D: Nonlinear Phenomena*, vol. 65, no. 1-2, pp. 86–99, 1993.

[95] L. E. Matson, "The malkus-lorenz water wheel revisited," *American Journal of Physics*, vol. 75, no. 12, pp. 1114–1122, 2007.

[96] E. N. Lorenz, "Deterministic Nonperiodic Flow.," *Journal of Atmospheric Sciences*, vol. 20, no. 2, pp. 130–148, 1963.

[97] J. La, G. Chen, D. Cheng, S. Celikovsky, and J, "Bridge the gap between the lorenz system and the chen system," *International Journal of Bifurcation and Chaos*, vol. 12, no. 12, pp. 2917–2926, 2002.

[98] J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, and E. Ott, "Using machine learning to replicate chaotic attractors and calculate lyapunov exponents from data," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, no. 12, p. 121102, 2017.

[99] P. Dubois, T. Gomez, L. Planckaert, and L. Perret, "Data-driven predictions of the lorenz system," *Physica D: Nonlinear Phenomena*, vol. 408, p. 132495, 2020.

[100] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, "Data-driven discovery of coordinates and governing equations," *Proceedings of the National Academy of Sciences*, vol. 116, no. 45, pp. 22445–22451, 2019.

[101] C.-I. Um, K.-H. Yeon, and T. F. George, "The quantum damped harmonic oscillator," *Physics Reports*, vol. 362, no. 2, pp. 63–192, 2002.

[102] M. Raissi, P. Perdikaris, and G. Karniadakis, "Multistep neural networks for data-driven discovery of nonlinear dynamical systems," *arXiv: 1801.01236*, 01 2018.

[103] A. Mendible, J. Koch, H. Lange, S. L. Brunton, and J. N. Kutz, "Data-driven modeling of rotating detonation waves," *Phys. Rev. Fluids*, vol. 6, p. 050507, May 2021.

[104] K. Hara and S. Cho, "Development of a hybrid particle-continuum kinetic method for hall thruster discharge plasmas," in *52nd AIAA/SAE/ASEE Joint Propulsion Conference*, (Salt Lake City, UT), AIAA-2016-4621, 2016.

[105] M. P. Georgin, B. A. Jorns, and A. D. Gallimore, "Correlation of ion acoustic turbulence with self-organization in a low-temperature plasma," *Physics of Plasmas*, vol. 26, no. 8,

p. 082308, 2019.

[106] R. Kawashima, K. Hara, and K. Komurasaki, "Numerical analysis of azimuthal rotating spokes in a crossed-field discharge plasma," *Plasma Sources Science and Technology*, vol. 27, no. 3, p. 035010, 2018.

[107] J. Fife, M. Martinez-Sanchez, and J. Szabo, "A numerical study of low-frequency discharge oscillations in hall thrusters," in *33rd Joint Propulsion Conference and Exhibit*, pp. AIAA–1997–3052, 1997.

[108] F. I. Parra, E. Ahedo, J. M. Fife, and M. Martínez-Sánchez, "A two-dimensional hybrid model of the hall thruster discharge," *Journal of Applied Physics*, vol. 100, no. 2, p. 023304, 2006.

[109] J. D. Adam, A. Héron, and G. Laval, "Study of stationary plasma thrusters using two-dimensional fully kinetic simulation," *Physics of Plasmas*, vol. 11, no. 1, pp. 295–305, 2004.

[110] D. Maddaloni, A. Domínguez-Vázquez, M. Merino-Martinez, and F. Terragni, "Data-driven analysis of breathing mode and ion-transit mode in 2d hybrid hall thruster simulations," in *AIAA Propulsion and Energy 2021 Forum*, (Virtual Events), 2021.

[111] M. Mörtl, A. Knoll, V. Williams, P. Shaw, V. Argyriou, J. Zamattio, and L. Pugliese, "Enhancing hall effect thruster simulations with deep recurrent networks," in *36th International Electric Propulsion Conference*, (Vienna, Austria), pp. IEPC–2019–501, 09 2019.

[112] K. Hara, M. J. Sekerak, I. D. Boyd, and A. D. Gallimore, "Perturbation analysis of ionization oscillations in hall effect thrusters," *Physics of Plasmas*, vol. 21, no. 12, p. 122103, 2014.

[113] A. J. Lotka, "Elements of physical biology," *Science Progress in the Twentieth Century (1919-1933)*, vol. 21, no. 82, pp. 341–343, 1926.

[114] V. Volterra, "Fluctuations in the abundance of a species considered mathematically1," *Nature*, vol. 118, no. 2972, pp. 558–560, 1926.

[115] E. T. Dale, B. Jorns, and K. Hara, "Numerical investigation of the stability criteria for the breathing mode in hall effect thrusters," *35th International Electric Propulsion Conference*,

pp. IEPC–2017–265, 2017.

[116] E. T. Dale and B. A. Jorns, "Non-invasive time-resolved measurements of anomalous collision frequency in a hall thruster," *Physics of Plasmas*, vol. 26, no. 1, p. 013516, 2019.

[117] 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," *IEEE Transactions on Plasma Science*, vol. 43, no. 1, pp. 72–85, 2015.

[118] G. N. Tilinin, "High-frequency plasma waves in a hall accelerator with an extended acceleration zone," *Sov. Phys. Tech. Phys.*, vol. 22, pp. 974–978, 1977.

[119] W. Huang, A. Gallimore, and T. Smith, "Laser-induced fluorescence of singly-charged xenon inside a 6-kw hall thruster," in *45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, (Denver, CO), pp. AIAA–2009–5355, 2009.

[120] C. M. Greve, M. Majji, and K. Hara, "Real-time state estimation of low-frequency plasma oscillations in hall effect thrusters," *Physics of Plasmas*, vol. 28, no. 9, p. 093509, 2021.

[121] M. Hayashi, "Determination of electron-xenon total excitation cross-sections, from threshold to 100 ev, from experimental values of townsend's α," *Journal of Physics D: Applied Physics*, vol. 16, no. 4, p. 581, 1983.

[122] V. Puech and L. Torchin, "Collision cross sections and electron swarm parameters in argon," *Journal of Physics D: Applied Physics*, vol. 19, no. 12, pp. 2309–2323, 1986.

[123] K. Tachibana, "Excitation of the $1s_5$,$1s_4$, $1s_3$, and $1s_2$ levels of argon by low-energy electrons," *Phys. Rev. A*, vol. 34, pp. 1007–1015, Aug 1986.

[124] S. Ashida, C. Lee, and M. A. Lieberman, "Spatially averaged (global) model of time modulated high density argon plasmas," *Journal of Vacuum Science & Technology A*, vol. 13, no. 5, pp. 2498–2507, 1995.

[125] V. A. Godyak and V. N. Maximov *Vestnik Moscovskogo Universiteta, Fisika i Astronomia*, vol. 51, no. 5, 1977.

[126] G. S. Hurst, E. B. Wagner, and M. G. Payne, "Energy transfer from the resonance states ar(^{1}p$_1$) and ar(^{3}p$_1$) to ethylene," *The Journal of Chemical Physics*, vol. 61, no. 9, pp. 3680–

3685, 1974.

[127] T. G. Owano, C. H. Kruger, and R. A. Beddini, "Electron-ion three-body recombination coefficient of argon," *AIAA Journal*, vol. 31, no. 1, pp. 75–82, 1993.

[128] D. Benoy, J. van der Mullen, M. van de Sanden, B. van der Sijde, and D. Schram, "Application of a hybrid collisional radiative model to recombining argon plasmas," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 49, no. 2, pp. 129–139, 1993.

[129] C. M. Ferreira and J. Loureiro, "Electron transport parameters and excitation rates in argon," *Journal of Physics D: Applied Physics*, vol. 16, no. 9, pp. 1611–1621, 1983.

[130] C. Lee, D. B. Graves, M. A. Lieberman, and D. W. Hess, "Global model of plasma chemistry in a high density oxygen discharge," *Journal of the Electrochemical Society*, vol. 141, no. 6, pp. 1546–1555, 1994.

[131] O. Kononchuk and B.-Y. Nguyen, *Silicon-on-insulator (soi) technology: Manufacture and applications*. Elsevier, 2014.

[132] D. P. Lymberopoulos, V. I. Kolobov, and D. J. Economou, "Fluid simulation of a pulsed-power inductively coupled argon plasma," *Journal of Vacuum Science & Technology A*, vol. 16, no. 2, pp. 564–571, 1998.

[133] S. Ashida, M. R. Shim, and M. A. Lieberman, "Measurements of pulsed-power modulated argon plasmas in an inductively coupled plasma source," *Journal of Vacuum Science & Technology A*, vol. 14, no. 2, pp. 391–397, 1996.

[134] J. Han, P. Pribyl, W. Gekelman, A. Paterson, S. J. Lanham, C. Qu, and M. J. Kushner, "Three-dimensional measurements of plasma parameters in an inductively coupled plasma processing chamber," *Physics of Plasmas*, vol. 26, no. 10, p. 103503, 2019.

[135] C. Qu, S. J. Lanham, S. C. Shannon, S. K. Nam, and M. J. Kushner, "Power matching to pulsed inductively coupled plasmas," *Journal of Applied Physics*, vol. 127, no. 13, p. 133302, 2020.

[136] J.-H. Park, D.-H. Kim, Y.-S. Kim, and C.-W. Chung, "Time-resolved spatial distribution measurements of pulse-modulated argon plasmas in an inductively coupled plasma reactor,"

Plasma Sources Science and Technology, vol. 26, no. 5, p. 055016, 2017.

[137] G. Park, H. Lee, G. Kim, and J. K. Lee, "Global model of he/o_2 and ar/o_2 atmospheric pressure glow discharges," *Plasma Processes and Polymers*, vol. 5, no. 6, pp. 569–576, 2008.

[138] J. T. Gudmundsson and E. G. Thorsteinsson, "Oxygen discharges diluted with argon: dissociation processes," *Plasma Sources Science and Technology*, vol. 16, pp. 399–412, 2007.

[139] G. Roberson, M. Roberto, J. Verboncoeur, and P. Verdonck, "Global model simulations of low-pressure oxygen discharges," *Brazilian journal of physics*, vol. 37, pp. 457–465, 2007.

[140] Meichsner, Jürgen and Wegner, Thomas, "Evaluation of oxygen species during e-h transition in inductively coupled rf plasmas: combination of experimental results with global model," *Eur. Phys. J. D*, vol. 72, no. 5, p. 85, 2018.

[141] W. Liu, D.-Q. Wen, S.-X. Zhao, F. Gao, and Y.-N. Wang, "Characterization of o_2/ar inductively coupled plasma studied by using a langmuir probe and global model," *Plasma Sources Science and Technology*, vol. 24, p. 025035, apr 2015.

[142] A. V. Vasenkov, X. Li, G. S. Oehrlein, and M. J. Kushner, "Properties of c-c_4f_8 inductively coupled plasmas. ii. plasma chemistry and reaction mechanism for modeling of ar/c-c_4f_8/o_2 discharges," *Journal of Vacuum Science & Technology A*, vol. 22, no. 3, pp. 511–530, 2004.

[143] S. Rauf, "On uniformity and non-local transport in low pressure capacitively coupled plasmas," *Plasma Sources Science and Technology*, vol. 29, no. 9, p. 095019, 2020.

[144] M. Schaepkens, R. C. M. Bosch, T. E. F. M. Standaert, G. S. Oehrlein, and J. M. Cook, "Influence of reactor wall conditions on etch processes in inductively coupled fluorocarbon plasmas," *Journal of Vacuum Science & Technology A*, vol. 16, no. 4, pp. 2099–2107, 1998.

[145] B. Zhou, E. A. Joseph, S. P. Sant, Y. Liu, A. Radhakrishnan, L. J. Overzet, and M. J. Goeckner, "Effect of surface temperature on plasma-surface interactions in an inductively coupled modified gaseous electronics conference reactor," *Journal of Vacuum Science & Technology A*, vol. 23, no. 6, pp. 1657–1667, 2005.

[146] K. J. Taylor and G. R. Tynan, "Control of dissociation by varying oxygen pressure in noble

gas admixtures for plasma processing," *Journal of Vacuum Science & Technology A*, vol. 23, no. 4, pp. 643–650, 2005.

[147] M. Cai, D. A. Haydar, A. Montaser, and J. Mostaghimi, "Computer simulation of argon-nitrogen and argon-oxygen inductively coupled plasmas," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 52, no. 3, pp. 369–386, 1997.

[148] M. Tadokoro, H. Hirata, N. Nakano, Z. L. Petrović, and T. Makabe, "Time resolved optical emission spectroscopy of an inductively coupled plasma in argon and oxygen," *Phys. Rev. E*, vol. 57, pp. R43–R46, Jan 1998.

[149] T. Sato and T. Makabe, "A numerical investigation of atomic oxygen density in an inductively coupled plasma in o_2/ar mixture," *Journal of Physics D: Applied Physics*, vol. 41, p. 035211, jan 2008.

[150] G. Evensen, "Using the extended kalman filter with a multilayer quasi-geostrophic ocean model," *Journal of Geophysical Research: Oceans*, vol. 97, no. C11, pp. 17905–17924, 1992.

[151] A. Yanou, N. Hosoya, K. Wada, M. Minami, and T. Matsuno, "Estimation of thermal conductivity for model with radiative heat transfer by extended kalman filter," in *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–4, 2016.

[152] A. Morozov, "The conceptual development of stationary plasma thrusters," *Plasma Physics Reports*, vol. 29, no. 3, pp. 235–250, 2003.

[153] A. I. Morozov, Y. V. Esinchuk, and G. N. Tilinin, "Plasma accelerator with closed electron drift and extended acceleration zone," *Sov. Phys.-Tech. Phys. (Engl. Transl.)*, vol. 17, no. 1, pp. 38–45, 1972.

[154] C. A. Lentz, *Transient one dimensional numerical simulation of Hall thrusters*. PhD thesis, MIT, 1993.

[155] W. A. Hargus Jr. and M. A. Cappelli, "Laser-induced fluorescence measurements of velocity within a hall discharge," *Applied Physics B*, vol. 72, no. 8, pp. 961–969, 2001.

[156] V. A. Fassel and R. N. Kniseley, "Inductively coupled plasma. optical emission spectroscopy," *Analytical Chemistry*, vol. 46, no. 13, pp. 1110A–1120a, 1974.

[157] K. Hara and I. G. Mikellides, "Characterization of low frequency ionization oscillations in hall thrusters using a one-dimensional fluid model," in *2018 Joint Propulsion Conference*, (Cincinnati, OH), pp. AIAA–2018–4904, 2018.

[158] J. T. Yim, *Computational Modeling of Hall Thruster Channel Wall Erosion*. PhD thesis, University of Michigan, 2008.

[159] K. Sawlani, *Effects of Secondary Electron Emission on the Electron Energy Distribution and Sheaths in Crossed Field Plasma Devices*. PhD thesis, University of Michigan, 2015.

[160] R. R. Hofer, D. M. Goebel, I. G. Mikellides, and I. Katz, "Magnetic shielding of a laboratory hall thruster. ii. experiments," *Journal of Applied Physics*, vol. 115, no. 4, p. 043304, 2014.

[161] C. Boniface, L. Garrigues, G. Hagelaar, J. Boeuf, D. Gawron, and S. Mazouffre, "Anomalous cross field electron transport in a hall effect thruster," *Applied Physics Letters*, vol. 89, no. 16, p. 161503, 2006.

[162] G. Hagelaar, J. Bareilles, L. Garrigues, and J.-P. Boeuf, "Role of anomalous electron transport in a stationary plasma thruster simulation," *Journal of Applied Physics*, vol. 93, no. 1, pp. 67–75, 2003.

[163] D. Byers and J. W. Dankanich, "A review of facility effects on hall effect thrusters," in *31st International Electric Propulsion Conference*, (Ann Arbor, MI), pp. AIAA–2009–076, 2009.

[164] J. D. Frieman, H. Kamhawi, G. Williams, W. Huang, D. A. Herman, P. Y. Peterson, J. H. Gilland, and R. R. Hofer, "Long duration wear test of the nasa hermes hall thruster," in *2018 Joint Propulsion Conference*, (Cincinnati, OH), pp. AIAA–2018–4645, 2018.

[165] J. D. Frieman, S. T. King, M. L. R. Walker, V. Khayms, and D. King, "Role of a conducting vacuum chamber in the hall effect thruster electrical circuit," *Journal of Propulsion and Power*, vol. 30, pp. 1471–1479, 2017/04/26 2014.

[166] K. D. Diamant, R. Liang, and R. L. Corey, "The effect of background pressure on spt-100

hall thruster performance," in *50th AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, (Cleveland, OH), pp. AIAA–2014–3710, 2014.

[167] K. Hara, I. D. Kaganovich, and E. A. Startsev, "Generation of forerunner electron beam during interaction of ion beam pulse with plasma," *Physics of Plasmas*, vol. 25, no. 1, p. 011609, 2018.

[168] G. J. M. Hagelaar and L. C. Pitchford, "Solving the boltzmann equation to obtain electron transport coefficients and rate coefficients for fluid models," *Plasma Sources Science and Technology*, vol. 14, no. 4, p. 722, 2005.

[169] D. E. Troyetsky, C. M. Greve, S. Tsikata, and K. Hara, "Estimation of the dynamic behavior of plasma properties in a hall effect thruster discharge plasma," *Submitted*, 2022.

[170] P. Y. Peterson, H. Kamhawi, W. Huang, G. Williams, J. H. Gilland, J. Yim, R. R. Hofer, and D. A. Herman, "Nasa's hermes hall thruster electrical configuration characterization," in *52nd AIAA/SAE/ASEE Joint Propulsion Conference*, (Salt Lake City, UT), pp. AIAA–2016–5027.

APPENDIX A

CURVE FITTING XENON RATE COEFFICIENTS

The xenon reaction rates can be found as tabulated data the BOLSIG+ database [168]. A fitting function was developed in MATLAB to fit resulting curves based on the functional forms presented in this section.

Using the equational form for the excitation reaction rate coefficient from Ref. 4, the radiative excitation and ** excitation reaction rates can be determined. Using the general equation form,

$$a\frac{\exp(b/T_e)}{\sqrt{T_e}}\sqrt{\frac{8eT_e}{\pi m_e}}, \tag{A.1}$$

the resulting coefficients from the curve fitting process are shown in Table A.1.

Table A.1: The curve fit coefficients for excitation reaction rates based on the excitation equation given in Ref. 4.

Reaction	a	b	Order of Magnitude
Radiative Excitation	41.0105	-13.0080	10^{-21}
** Excitation	81.6125	-8.4280	10^{-21}

The equational form for ionization is written as

$$(a + bT_e + cT_e^2)\exp(d/T_e)\sqrt{\frac{8eT_e}{\pi m_e}}, \tag{A.2}$$

based on the equation provided in Ref. 4. This equation is used to fit the direct ionization rate coefficient data using the coefficients provided in Table A.2.

The final functional form used is based on the ionization reaction rate equation used above, but includes an exponential term such that the form becomes

Table A.2: The curve fit coefficients for the direct ionization reaction rate based on the ionization equation given in Ref. 4.

a	b	c	d	Order of Magnitude
6.9383	-0.03014	6.7079×10^{-5}	-12.3028	10^{-20}

$$(a + bT_e^c + \frac{d}{T_e} + e\exp(-\frac{T_e}{y\Delta\epsilon}))\sqrt{\frac{8e_{ch}T_e}{pim_e}} \tag{A.3}$$

This form is used to create curve fits for both the stepwise ionization and the metastable excitation reaction rates. The resulting coefficients are presented in Table A.3. Note that the coefficient e is different than the elementary charge of particles, e_{ch}.

Table A.3: The curve fit coefficients for stepwise ionization and metastable excitation reaction rates based on a modified ionization equation.

Reaction	a	b	c	d	e	g	Order of Magnitude
Stepwise Ionization	0.04192	-3.4567×10^{-9}	0.2803	-421.2596	-2.1090×10^{-8}	5.5788	10^{-21}
Metastable Excitation	0.002359	-1.6635×10^{-7}	1.4954×10^{-4}	-23.7056	-1.6747×10^{-10}	3.6978	10^{-21}