







RESEARCH ARTICLE | APRIL 17 2024

Bayesian plasma model selection for Thomson scattering

Jean Luis Suazo Betancourt  ; Samuel J. Grauer ; Junhwi Bak ; Adam M. Steinberg ;
Mitchell L. R. Walker 

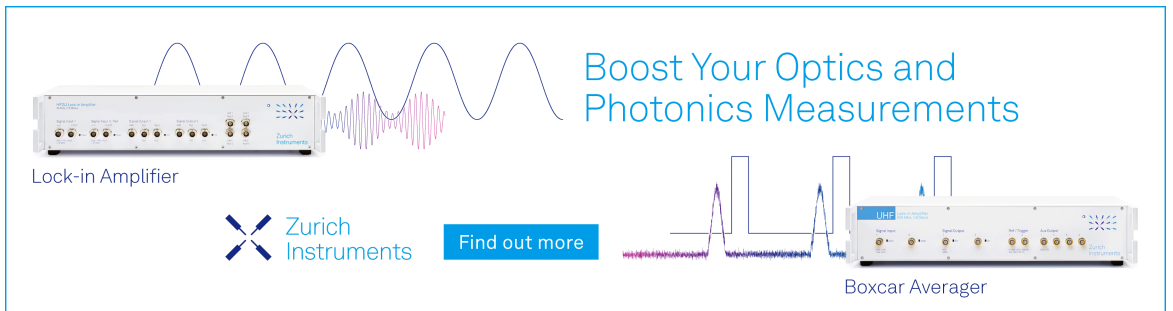
 Check for updates

Rev. Sci. Instrum. 95, 043004 (2024)

<https://doi.org/10.1063/5.0158749>



View
Online


Export
Citation



Boost Your Optics and Photonics Measurements

Lock-in Amplifier

 Zurich Instruments

[Find out more](#)

Boxcar Averager

Bayesian plasma model selection for Thomson scattering

Cite as: Rev. Sci. Instrum. 95, 043004 (2024); doi: 10.1063/5.0158749

Submitted: 17 May 2023 • Accepted: 24 March 2024 •

Published Online: 17 April 2024







View Online



Export Citation



CrossMark

Jean Luis Suazo Betancourt,^{1,a)}  Samuel J. Grauer,²  Junhwi Bak,³  Adam M. Steinberg,¹ 
and Mitchell L. R. Walker¹ 

AFFILIATIONS

¹School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, Georgia 30318, USA

²Department of Mechanical Engineering, Pennsylvania State University, University Park, Pennsylvania 16802, USA

³Department of Aerospace Engineering, Texas A&M University, College Station, Texas 77843, USA

^{a)}Author to whom correspondence should be addressed: jlsb3@gatech.edu

ABSTRACT

Laser Thomson scattering (LTS) is a measurement technique that can determine electron velocity distribution functions in plasma systems. However, accurately inferring quantities of interest from an LTS signal requires the selection of a plasma physics submodel, and comprehensive uncertainty quantification (UQ) is needed to interpret the results. Automated model selection, parameter estimation, and UQ are particularly challenging for low-density, low-temperature, potentially non-Maxwellian plasmas like those created in space electric propulsion devices. This paper applies Bayesian inference and model selection to a Raman-calibrated LTS diagnostic in the context of such plasmas. Synthetic data are used to explore the performance of the method across signal-to-noise ratios and model fidelity regimes. Plasmas with Maxwellian and non-Maxwellian velocity distributions are well characterized using priors that span a range of accuracy and specificity. The model selection framework is shown to accurately detect the type of plasmas generating the electron velocity distribution submodel for signal-to-noise ratios greater than around 5. In addition, the Bayesian framework validates the widespread use of 95% confidence intervals from least-squares inversion as a conservative estimate of the uncertainty bounds. However, epistemic posterior correlations between the variables diverge between least-squares and Bayesian estimates as the number of variable parameters increases. This divergence demonstrates the need for Bayesian inference in cases where accurate correlations between electron parameters are necessary. Bayesian model selection is then applied to experimental Thomson scattering data collected in a nanosecond pulsed plasma, generated with a discharge voltage of 5 and 10 kV at a neutral argon background pressure of 7 Torr-Ar. The Bayesian maximum *a posteriori* estimates of the electron temperature and number density are 1.98 and 2.38 eV and 2.6×10^{18} and $2.72 \times 10^{18} \text{ m}^{-3}$, using the Maxwellian and Druyvesteyn submodels, respectively. Furthermore, for this dataset, the model selection criterion indicates strong support for the Maxwellian distribution at 10 kV discharge voltage and no strong preference between Maxwellian and Druyvesteyn distributions at 5 kV. The logarithmic Bayes' factors for these cases are -35.76 and 1.07 , respectively.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0158749>

I. INTRODUCTION

This paper presents a Bayesian framework for plasma model selection and uncertainty quantification (UQ) in the context of laser Thomson scattering (LTS) measurements in low-density, low-temperature, potentially non-equilibrium plasma. While particularly motivated by plasmas in space electric propulsion (EP) thrusters, the framework is equally applicable to other LTS plasma measurement scenarios. Plasmas in EP thrusters are generated and

manipulated by electromagnetic fields to accelerate ions for thrust and to confine electrons for efficient ionization. Electrostatic EP devices, such as Hall effect thrusters (HETs), generate thrust by accelerating ions in an electrostatic field. HETs specifically use electrons confined in a radial magnetic field to ionize the neutral propellant, which is then accelerated in an axial electrostatic field. Many aspects of HET performance are difficult to model, and minimally invasive laser-based measurements of HET plasmas are needed to baseline the plasma simulations used to design new thrusters. Recent

laser-based studies of HETs indicate that the signal-to-noise ratio (SNR) is a challenge¹⁻³ that may necessitate robust signal inversion methods.

Electron dynamics have a significant influence on plasma behaviors due to the high mobility of electrons relative to ions.⁴ As a result, electron properties can be used to infer plasma phenomena in HETs. For instance, the electron temperature and density may be used to inform anomalous electron mobility, magnetic confinement and shielding, sheath size estimates at non-plasma boundaries, and the onset and extent of ionization and acceleration zones. However, accurate characterization of electron properties in HETs remains an open problem due to challenges in both acquiring and interpreting quality data.

Models that depend on electron motion are typically taken to be a function of Maxwellian thermal plasma properties. This is because making measurements of more general quantities like the electron velocity distribution function (EVDF) in certain plasma environments of interest, such as the near-field plume of HETs, is challenging using electrostatic probes. The harsh environment oftentimes destroys the probe or requires very short residence times. Furthermore, the probe affects the local plasma environment, bringing estimates of the plasma's (presumed) equilibrium state into question. Several EVDF models exist with which to infer the summary information like the equilibrium or effective electron temperature, and the most accurate model is not known *a priori*.

Giono *et al.*⁵ used a cylindrical Langmuir probe to take far-field measurements (at greater than ten thruster diameters) of electron energy probability functions (EPPFs) in a mid-power, traditionally unshielded SPT-100 HET.⁶ The authors concluded that the distribution functions were best represented by a mixed-model Maxwellian–Druyvesteyn EEPF, with the non-equilibrium EEPFs in the plume being linked to the thruster mass flow rate. The results in this paper indicated the dependencies of the plume EVDF on the discharge channel EVDF and collision rate. The high-power unshielded HET measurements with internal and external reverse-orientation cathodes have indicated both (i) Druyvesteyn EVDFs in the plume and (ii) the dependence of the EVDF on the cathode's orientation and placement.⁷ Furthermore, laser-induced fluorescence measurements in the near-field plume of the H9 magnetically shielded HET show skewed and non-Gaussian ion velocity distribution functions, suggesting the presence of non-Maxwellian, multi-distribution EVDFs.⁸ Given these HET plume EVDF results, as well as the need to validate models for the predictive simulation of electron properties in the near-field plume of magnetically shielded HETs, robust EVDF model selection and electron property estimates are needed.

LTS is an optical diagnostic that can characterize HET-relevant plasmas. LTS yields 1D EVDFs, from which the electron temperature, number density, bulk drift velocity, and other important properties can be determined. However, estimates of these quantities of interest (QoI) are contingent on plasma physics submodels, of which the applicability is not always known *a priori* for a given plasma. In addition, the results are often very sensitive to noise. For this reason, converting an LTS signal to physical QoI is particularly challenging in low-temperature, low-density plasmas, such as those found in HETs. Model selection and comprehensive UQ are thus needed to baseline plasma models using LTS data. Here,

we present and evaluate a Bayesian framework for LTS model selection and signal inversion with UQ, motivated by applications under plasma conditions relevant to HETs. However, the framework is applicable to signals collected from LTS in all plasmas, including non-Maxwellian and low-density plasmas.

LTS measurements have been successfully carried out in systems ranging from quiescent DC discharge plasmas to non-equilibrium pulsed plasmas.⁹ More recently, LTS has been leveraged to study plasma parameters in the plume of standard and magnetically shielded HETs.³ Several studies assessed uncertainties in these measurements, most commonly using least-squares sensitivities,¹⁰ with select examples of maximum likelihood estimation¹¹ and full Bayesian inference.¹²⁻¹⁵ It is noteworthy that these LTS UQ frameworks were developed for fusion-based applications. However, the detection schemes, equipment, and sources of uncertainty vary between the LTS systems used in fusion research and those used to measure EP devices. Fusion applications, necessitated by the high temperature and density environment, use several unintensified detectors and wavelength filters to capture intensity data.^{14,15} Recently implemented LTS systems for low-temperature and -density plasmas used a single intensified spectrograph to capture intensity data at multiple wavelengths simultaneously. In the fusion case, the acquisition of a spectrum requires the aggregation of intensity data from multiple detectors that can lead to wavelength uncertainty. In the EP relevant case, spectrum data can be captured directly with higher wavelength accuracy but oftentimes marred by low SNRs.

The primary challenge when measuring an HET plasma with LTS is the low SNR. HET plasmas typically exhibit a low number density (n_e from 10^{16} to 10^{20} m^{-3}) and large temperature ranges (T_e from 0.1 to upward of 50 eV) compared to other plasmas, such as arc discharge or fusion plasmas, whose number densities and temperatures can be orders of magnitude higher. Low electron densities in HET plasmas lead to less scattered radiation than in other systems, reducing the signal strength and thereby lowering the SNR. Signal throughput limitations in a typical detection system for HET applications also weaken the signal. Due to these factors, it is usually necessary to acquire hundreds to thousands of individual spectra to obtain a stable mean spectrum; processing workflows often assume a statistically stationary target and neglect the effects of nonlinearities in the measurement model. Several other sources of uncertainty can affect the interpretation of LTS measurements. For example, estimating n_e via LTS requires calibration with a secondary scattering technique, typically Rayleigh or rotational Raman scattering, a.k.a. laser Raman scattering (LRS), which are themselves subject to uncertainties that propagate through the parameter estimation workflow.

Plasma parameters are estimated from LTS data by inverting a forward measurement model that maps the (unknown) QoI to an LTS signal. In the simplest case of one LTS plasma submodel and one electron number density calibrating model, such as LRS, the parameter space of the highly nonlinear models is large, having anywhere from 5 to 15 unknown parameters. There exist several equilibrium and non-equilibrium candidate submodels that can have different parameterizations and result in different computed QoI. The signal analysis procedure is said to be well-posed when the signal and chosen forward model correspond to a unique solution that is robust

to noise, in which case a simple- or weighted-least-squares algorithm may be used to calculate the QoI. However, if the forward map is degenerate, meaning that multiple QoI yield the same or highly similar signals, then the problem is ill-posed, and additional “prior” information must be included to regularize the inversion. This supplemental information is itself a source of uncertainty, which must be accounted for in the inversion scheme to get an accurate estimate of the parameter uncertainties.

Bayesian inference has been employed in several EP applications, such as the analysis of triple Langmuir probe data,¹⁶ terahertz time-domain spectroscopy,¹⁷ and EP-based model inference for electrospray thrusters.¹⁸ To address the aforementioned signal inversion issues, this paper introduces a Bayesian framework for parameter estimation, UQ, and model selection in LTS experiments calibrated by Raman scattering. Synthetic data are used to quantify the SNR regimes in which a distinction can be made between candidate submodels for Maxwellian and non-Maxwellian EVDFs. We show that the Laplace approximation to the posterior is appropriate at moderate SNRs, giving accurate estimates of the QoI in LRS-calibrated LTS experiments. In addition, we show that for moderate SNRs and above, the model selection framework can accurately distinguish electron velocity distribution submodels. Finally, we analyze experimental LTS measurements of a nanosecond-pulsed plasma.

II. LASER RAMAN AND THOMSON SCATTERING MODELS

Thomson scattering is the elastic scattering of incident radiation from unbounded electrons, which is spectrally broadened and shifted due to random and bulk motion of the electrons, respectively. In effect, LTS provides a direct measurement of the EVDF and bulk electron velocity. Furthermore, given a proper system calibration, the amplitude of the LTS spectrum also indicates the electron number density. The electron number density is a critical state parameter in HET applications, making the intensity calibration necessary.

LRS is the preferred technique for calibrating LTS because Raman scattering is less susceptible to interference at the laser wavelength, e.g., due to surface reflections, than alternatives such as Rayleigh scattering.¹⁹ By contrast, Raman scattering is inelastic, so interference from the laser can be mitigated with a filter.

Forward models for both Raman and Thomson scattering, indicated by “R” and “T” superscripts in this paper, respectively, feature a large number of parameters. Some of these parameters are deemed QoI, or “quantities of interest,” denoted by the vector \mathbf{x} , and others are “nuisance parameters,” i.e., unknown variables that are not of primary interest, denoted $\boldsymbol{\theta}$. Nuisance parameters, such as the collection efficiency of an LTS system, must be considered since they are not known *a priori* and are subject to uncertainties that affect the estimates of the QoI. In other words, it is important to account for uncertainties about $\boldsymbol{\theta}$ when estimating \mathbf{x} , conducting UQ, and comparing system models.

The LRS intensity in units of counts per nm at wavelength λ , produced by monochromatic incident light at wavelength λ_i and intensity E_i , is

$$P_{\lambda}^R(\mathbf{x}^R, \boldsymbol{\theta}^R) = \eta \frac{\lambda_i}{hc} \Delta\Omega L E_i n_g(T_g, p_g) \sum_j \frac{\partial \sigma^R}{\partial \Omega_j} S_{\lambda}^R(\lambda_j, \tau), \quad (1)$$

where η is the collection efficiency of the optical system, $\Delta\Omega$ is the solid angle of collection, L is the length of the probe volume, and λ_i/hc is the conversion from incident energy to photon counts. The variables J and J' represent the population levels, as described in [Appendix A 1](#). L and $\Delta\Omega$ are estimated using the beam waist image size on the spectrometer slit plane, scaled by the detection system’s magnification, and a thin lens optical analysis of the collection system optics. In this way, $\Delta\Omega$ is used as an optimization parameter for the collection and detection systems, and η absorbs all forms of system inefficiencies.

The LRS spectrum depends upon the neutral gas temperature, T_g , and has an intensity proportional to the incident laser energy, E_i , and gas number density, n_g . These parameters, in turn, depend on T_g and the pressure, p_g . It is generally necessary to measure p_g , as we have done in the experiment outlined in [Sec. V](#), whereas T_g can be measured, inferred, or assumed to be a constant, e.g., room temperature, as is often done. In cases where T_g is a variable of interest or unmeasured, as in our experiments, it can be inferred from an LRS spectrum so long as p_g is independently measured.²⁰ In the case where neither is measured, the model can be classified as non-identifiable, given that η and p_g have the same effect on the spectra. This non-identifiability can be mitigated through strong prior information. In addition, it may be possible to get a unique estimate for η even if p_g and T_g are unknown using Bayesian inference methods and MCMC sampling. However, these methods would be inefficient. The width of broadened Raman lines is in large part dependent on the instrument broadening under standard calibration conditions, where other broadening mechanisms are negligible. This is typically expressed in terms of the lines’ full width at half maximum (FWHM). We impute a single FWHM to all the Raman lines, denoted τ .

The differential Raman scattering cross section of the gas is $\partial \sigma^R / \partial \Omega$, and S_{λ}^R is the spectral distribution function. Note that all parameters in the LRS model are summarized in [Table I](#), and the details of the Raman sub-functions are presented in [Appendix A 1](#). The Raman QoI and nuisance parameters are

$$\mathbf{x}^R = [\eta, \lambda_i]^T \quad \text{and} \quad \boldsymbol{\theta}^R = [T_g, p_g, \tau]^T. \quad (2)$$

Given known laser, gas, and spectral parameters, the measurement of P_{λ}^R from neutral gases can be used to determine η and λ_i for subsequent use in LTS measurements. Therefore, these QoI from the Raman inference are nuisance parameters in the Thomson inference. [Figure 1\(a\)](#) depicts a Raman spectrum that was simulated using the sample parameters listed in [Table I](#).

The Thomson scattering signals are described by

$$P_{\lambda}^T(\mathbf{x}^T, \boldsymbol{\theta}^T) = \eta \frac{\lambda_i}{hc} \Delta\Omega L E_i n_e \frac{\partial \sigma^T}{\partial \Omega} S_{k,\lambda}^T(\mathbf{x}^T, \lambda_i), \quad (3)$$

where $\partial \sigma^T / \partial \Omega$ is the Thomson scattering cross section and η is the system constant obtained from LRS calibration. Notably, $S_{k,\lambda}^T$ is a general spectral distribution function for Thomson scattering, which depends on the plasma submodel and is parameterized by a vector of model-specific variables, \mathbf{x}^T . We model the target plasma as a superposition of Maxwellian and Druyvesteyn distributions, each

TABLE I. Units and descriptions for the parameters used in the LRS models presented in Sec. II and Appendix A 1. The values in this table were used to generate the synthetic spectrum presented in 1a.

Parameter	Unit	Value	Description
p_g	Torr	10	Measured neutral gas pressure
T_g	K	300	Neutral gas temperature inferred from the LRS spectrum
τ	nm	0.153	LRS spectral linewidth FWHM, inferred from the LRS spectrum
η	...	1	System efficiency constant, inferred from the LRS spectrum
λ_i	nm	532	Incident laser wavelength, inferred from the LRS spectrum
E_i	mJ	200	Measured incident laser energy per pulse
γ^2/ϵ_0^2	m^6	0.505×10^{-60}	Polarizability tensor anisotropy from Ref. 21
ζ	rad	0	Raman polarization angle assumed to be parallel to the incident laser polarization
$\Delta\Omega$	sr	0.05	Calculated system solid angle
L	mm	2	Interrogation volume length, estimated from the beam waist image
θ_{eiy}	rad	0	Incident polarization to y angle, parallel to our lab-frame y
θ_{kio}	rad	$\pi/2$	Incident to collection k angle, perpendicular to k and parallel to our lab-frame z
ρ	...	3/4	Depolarization ratio from Ref. 21
I_g	...	1	Nuclear spin quantum number from Ref. 21
B_g	m^{-1}	198.973	Rotational energy constant from Ref. 21
D_g	m^{-1}	5.7×10^{-4}	Centrifugal distortion constant from Ref. 22
$g_{j=\text{even}}$...	6	Nuclear spin degeneracy from Ref. 21
$g_{j=\text{odd}}$...	3	Nuclear spin degeneracy from Ref. 21

specified in terms of a distinct electron number density, n_e , and temperature, T_e ,

$$S_{k,\lambda}^T(\mathbf{x}^T, \boldsymbol{\theta}^T) = \frac{1}{n_e} \left[\sum_{i=1}^{N_M} n_{e,i}^M S_{k,\lambda}^M(n_{e,i}^M, T_{e,i}^M) + \sum_{i=1}^{N_D} n_{e,i}^D S_{k,\lambda}^D(n_{e,i}^D, T_{e,i}^D) \right]. \quad (4)$$

In this expression, S_k^M and S_k^D are the Maxwellian and Druyvesteyn distribution functions, outlined in Appendix A 3; N_M and N_D denote the number of each distribution included in the mixture; and n_e is the total electron number density. All parameters of the Thomson scattering model are summarized in Table II. There is support for constructing distribution functions as weighted sums of analytical distribution functions in the plasma literature at large, with the most common being the bi-Maxwellian distribution; see Appendix A 3.^{23–28}

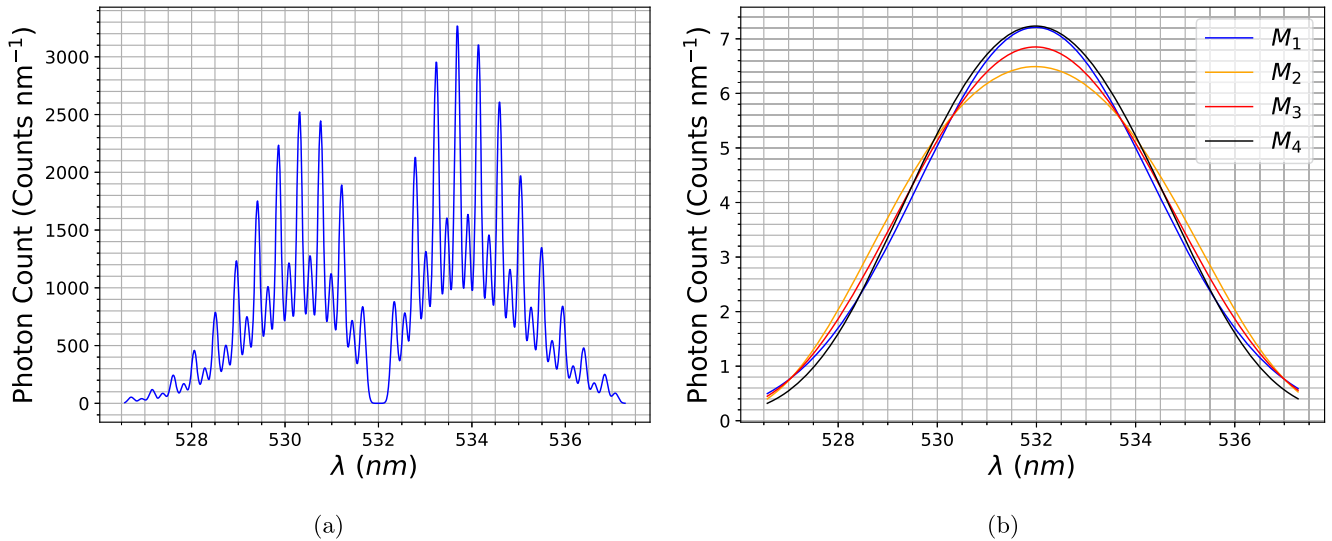

FIG. 1. Synthetic spectra generated with parameters from Tables I and II. (a) LRS spectrum and (b) LTS spectra for the submodels in Table III.

TABLE II. Units and descriptions for the parameters used in the LTS models presented in Table II, Appendix A 2, and Appendix A 3. The values in this table were used to generate the synthetic spectrum presented in 1b.

Parameter	Unit	Value	Description
γ_{MD}		0.5	Maxwellian–Druyvesteyn distribution fraction
n_e^M	m^{-3}	1×10^{17}	Maxwellian number density
n_e^D	m^{-3}	1×10^{17}	Druyvesteyn number density
T_e^M	eV	5	Maxwellian electron temperature
T_e^D	eV	5	Druyvesteyn electron temperature
v_d	$km\ s^{-1}$	0	Bulk drift velocity, inferred from the LTS spectrum
η		1	System efficiency constant, inferred from the LRS spectrum
λ_i	nm	532	Incident laser wavelength, inferred from the LRS spectrum
E_i	mJ	200	Measured incident laser energy per pulse
n		1	Refractive index
$\Delta\Omega$	sr	0.05	Calculated system solid angle
L	mm	2	Interrogation volume length, estimated from the beam waist image
θ_{eiy}	rad	0	Incident polarization to y angle, parallel to our lab-frame y
θ_{kio}	rad	$\pi/2$	Incident to collection k angle, perpendicular to k and parallel to our lab-frame z
r_e	m	2.8179×10^{-15}	Classical electron radius from Ref. 10

The QoI and nuisance parameter vectors in the Thomson model are

$$\mathbf{x}^T = [\mathbf{T}^M; \mathbf{n}^M; \mathbf{T}^D; \mathbf{n}^D; v_d]^T \quad \text{and} \quad \boldsymbol{\theta}^T = [\eta, \lambda_i]^T, \quad (5)$$

with \mathbf{n} and \mathbf{T} being the electron number density and temperature vectors, for instance,

$$\mathbf{n} = [n_{e,1}, n_{e,2}, \dots, n_{e,N_M}]^T. \quad (6)$$

The superscripts indicate whether a vector represents the Maxwellian or Druyvesteyn distribution. Using this notation, an equilibrium plasma corresponds to the model $N_M = 1$, $N_D = 0$, and $\mathbf{x}^T = [T_e, n_e, v_d]^T$. In this work, we limit the superposition to cases where both N_M and N_D are less than or equal to one. For convenience, we define a distribution fraction for binary Maxwellian–Druyvesteyn plasmas ($N_M = N_D = 1$),

$$\gamma^{MD} = \frac{n_e^M}{n_e^M + n_e^D} = \frac{n_e^M}{n_e}. \quad (7)$$

Restricting our model to a low-order sum is consistent with a generalized distribution commonly found in the literature; see Refs. 24 and 26–28. There is also support in the literature for extending this to higher-order sums (see Appendix A 3). Understanding the challenges and trade-offs in accurate model selection and signal inversion for higher-order sums is a topic of future research.

Four plasma models are considered in this work, as specified in Table III. Note that model M_3 features a binary Maxwellian–Druyvesteyn plasma with a constant electron temperature, i.e., $T_e = T_e^M = T_e^D$, whereas M_4 contains two distinct temperatures. For Thomson scattering, the QoI are given by Eq. (5), which differs depending on the model, defined by N_M , N_D , v_d , and the vector of nuisance parameters from the Raman inference. The simulated Thomson spectra based on M_1 – M_4 , using the parameters listed in Table II, are shown in Fig. 1(b).

The LRS and LTS models presented in Eqs. (1) and (3) produce an intensity at each wavelength, λ . A measured power spectrum can be represented in vector form,

$$\mathbf{b} = [P_{\lambda_1}, P_{\lambda_2}, \dots, P_{\lambda_{N_\lambda}}]^T, \quad (8)$$

where N_λ is the number of spectral resolution elements. It is often-times necessary to reject those parts of a spectrum that straddle the laser wavelength. Therefore, when evaluating experimental data, elements of \mathbf{b} within 4τ of the laser wavelength are eliminated from the comparison. In what follows, we use both real data, synthetic data and modeled data. The synthetic data are indicated with a tilde, e.g.,

$$\tilde{\mathbf{b}} = [\tilde{P}_{\lambda_1}, \tilde{P}_{\lambda_2}, \dots, \tilde{P}_{\lambda_{N_\lambda}}]^T, \quad (9)$$

and the modeled data are indicated with a hat, e.g.,

$$\hat{\mathbf{b}}(\mathbf{x}) = [P_{\lambda_1}(\mathbf{x}), P_{\lambda_2}(\mathbf{x}), \dots, P_{\lambda_{N_\lambda}}(\mathbf{x})]^T. \quad (10)$$

TABLE III. Description of the different plasma submodel QoI and nuisance parameter vectors used for Bayesian model selection. M_1 and M_2 describe the single Maxwellian and Druyvesteyn submodels. M_3 and M_4 describe the single and distinct temperature sums of Maxwellian and Druyvesteyn submodels.

Model	\mathbf{x}^T	$\boldsymbol{\theta}^T$
M_1	$[T_e^M, n_e^M, v_d]^T$	$[\eta, \lambda_i]^T$
M_2	$[T_e^D, n_e^D, v_d]^T$	$[\eta, \lambda_i]^T$
M_3	$[T_e, n_e, \gamma^{MD}, v_d]^T$	$[\eta, \lambda_i]^T$
M_4	$[T_e^M, T_e^D, n_e, \gamma^{MD}, v_d]^T$	$[\eta, \lambda_i]^T$

III. BAYESIAN INFERENCE AND MODEL SELECTION

A. Background and Bayes' equation

Bayesian inference is the process of updating a prior probability distribution given an observation.^{29,30} Often in optical experiments, the measurement models are high-fidelity, dependent on many parameters, and a single optical setup comprises multiple diagnostics that provide partial information about the QoI. In the context of parameter estimation and UQ, an uncertainty model is developed based on prior information about the QoI and then updated subject to the measurement data. The resultant “posterior” information may be employed as prior information in subsequent inferences and refined in light of new data. This approach can be useful in a staged inference problem, where QoI from an earlier inference are nuisance parameters in a subsequent inference; this is the case for Bayesian inference applied to LRS-calibrated LTS data. Bayesian inference has extensively been leveraged for UQ by the fusion community.³¹ However, LTS for EP applications entails unique considerations compared to fusion applications due to their distinctive pre-processing requirements and model equations. For instance, advances in detection, dispersion, and filtering technology have enabled the use of a single intensified-spectrograph system, which have sufficient spectral range and resolution to determine the QoI.¹⁹ This section describes a Bayesian framework for evaluating the LRS-calibrated LTS measurements of EP-relevant plasmas.

The QoI, measured signals, physical parameters, and models are cast as random variables that are characterized by probability distributions. These distributions describe one's knowledge of the parameters—narrow distributions indicate a high degree of confidence, while wide distributions signify uncertainty. The posterior distribution, $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta} | \mathbf{b})$,³² conveys one's knowledge of the QoI following a measurement. There are several distributions of interest in our framework. The likelihood, which is the PDF that determines how likely a given data \mathbf{b} was produced by a set of QoI and nuisance parameters. The prior PDF, or simply prior, $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta})$, is the joint PDF of the QoI and nuisance parameters based on any information about these variables that is independent of the measurement. Prior information may include physical limits, results from previous experiments, and so on. The posterior is the joint distribution of the QoI and nuisance parameters, conditioned on the observed data through the likelihood.

The posterior is a comprehensive Bayesian solution to an inverse problem since it carries all the measured and prior information about the QoI and nuisance parameters. The prior and likelihood are constructed to include as much information about the system and measurement scenario as is practical. The posterior, $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta} | \mathbf{b})$, likelihood, $\mathbb{P}(\mathbf{b} | \mathbf{x}, \boldsymbol{\theta})$, prior, $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta})$, and evidence, $\mathbb{P}(\mathbf{b})$, are related by Bayes' equation,

$$\mathbb{P}(\mathbf{x}, \boldsymbol{\theta} | \mathbf{b}) = \frac{\mathbb{P}(\mathbf{b} | \mathbf{x}, \boldsymbol{\theta}) \mathbb{P}(\mathbf{x}, \boldsymbol{\theta})}{\mathbb{P}(\mathbf{b})}. \quad (11)$$

The evidence itself, a.k.a. the marginal likelihood, is the marginal joint distribution,

$$\mathbb{P}(\mathbf{b}) = \iint \mathbb{P}(\mathbf{b} | \mathbf{x}, \boldsymbol{\theta}) \mathbb{P}(\mathbf{x}, \boldsymbol{\theta}) d\mathbf{x} d\boldsymbol{\theta}. \quad (12)$$

The posterior is summarized using point estimates, such as the maximum *a posteriori* (MAP) estimate of the QoI, \mathbf{x}_{MAP} . The MAP

point estimate is the estimate that maximizes the nuisance parameter marginalized posterior,

$$\mathbb{P}(\mathbf{x} | \mathbf{b}) = \iint \mathbb{P}(\mathbf{x}, \boldsymbol{\theta} | \mathbf{b}) d\boldsymbol{\theta}. \quad (13)$$

Point estimates are not sufficient for UQ, which requires at least the spread of the posterior about the MAP to convey an estimate of uncertainty.

B. Likelihood and prior PDFs

A measurement model that is a function of the QoI, $\hat{\mathbf{b}}(\mathbf{x}, \boldsymbol{\theta})$, is the basis for determining the QoI in any optical experiment. In this case, the model corresponds to either the Raman or Thomson model from Sec. II. These models are close approximations to the real physical process being observed, but discrepancies arise due to noise and other imperfections in the system. The likelihood function quantifies the chance that the data, \mathbf{b} , could have been observed given a hypothetical set of QoI and nuisance parameters, \mathbf{x} and $\boldsymbol{\theta}$. This probability corresponds to the distribution of errors,

$$\mathbf{e}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{b}) \equiv \mathbf{b} - \hat{\mathbf{b}}(\mathbf{x}, \boldsymbol{\theta}), \quad (14)$$

where \mathbf{x} and $\boldsymbol{\theta}$ are ideal parameters for a set of data.

In this work, the data are a spectrum, assumed to have been preprocessed (binned, etc.) ahead of being passed to the framework. See Chap. 2 of Ref. 10 for a detailed explanation of pre-processing of raw LTS and LRS signals. The starting point for inversion is a spectrum that is background- and reflection-corrected by subtracting the plasma emission spectrum and the background reflection spectrum from a raw LTS spectrum. Common sources of noise—such as thermal noise in detection systems, read noise on the face of a solid-state detector, and many others—can be approximated as Gaussian; see Refs. 32 and 33 for details. Moreover, macroscopic representations of measurement noise correspond to the macro-scale averaging of many noise-creating processes and, as such, adhere to the central limit theorem.³²

Therefore, we choose to cast the likelihood function in this work as a Gaussian likelihood, given by

$$\mathbb{P}(\mathbf{b} | \mathbf{x}, \boldsymbol{\theta}) = \det(2\pi\boldsymbol{\Gamma}_e)^{-1/2} \exp\left(-\frac{1}{2}\mathbf{e}^\top \boldsymbol{\Gamma}_e^{-1} \mathbf{e}\right), \quad (15)$$

with $\boldsymbol{\Gamma}_e \in \mathbf{R}^{d_b \times d_b}$ being the error covariance matrix. The size of the data, d_b —in this case the number of points making up a spectrum—is on the order of the number of pixels in a row on the spectrograph detector (usually in the thousands). Assuming that the errors are independent and identically distributed, with variance σ_e^2 , then $\boldsymbol{\Gamma}_e = \sigma_e^2 \mathbf{I}$ and the likelihood simplifies to

$$\mathbb{P}(\mathbf{b} | \mathbf{x}, \boldsymbol{\theta}) = (2\pi\sigma_e^2)^{-d_b/2} \exp\left(-\frac{\mathbf{e}^\top \mathbf{e}}{2\sigma_e^2}\right). \quad (16)$$

The prior PDF incorporates information about the QoI and nuisance parameters that is independent of the data. Any prior information that can be explicitly encoded in a probability distribution can be included. This information may include physical bounds, dynamical relations, previous estimates of the QoI, etc. We

assume that our prior information about the QoI is uncorrelated and Gaussian, in which case

$$\mathbb{P}(\mathbf{x}) = \prod_{i=1}^{d_x} \frac{1}{\sqrt{2\pi\sigma_{x,i}^2}} \exp\left[-\frac{(x_i - \mu_{x,i})^2}{2\sigma_{x,i}^2}\right], \quad (17)$$

where

$$\mu_{x,i} = \min(x_i) + \frac{\max(x_i) - \min(x_i)}{2} \quad (18)$$

is the center of the physically plausible range of values for x_i and

$$\sigma_{x,i} = \frac{1}{3} [\max(x_i) - \min(x_i)] \quad (19)$$

ensures that 99.7% of the PDF occupies this range. The same procedure is used to define $\mathbb{P}(\boldsymbol{\theta})$, and since the QoI and nuisance parameters are independent, we set $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta})$ to $\mathbb{P}(\mathbf{x})\mathbb{P}(\boldsymbol{\theta})$. The variables d_x and d_θ represent the length of the QoI and nuisance parameter vectors, respectively.

C. Sampling from and approximating the posterior PDF

The posterior PDF is a high-dimensional function that is not generally available in closed form. We consider two methods for accessing the posterior for UQ and model selection: sampling and Laplace's approximation. Given the dimension of the data, d_b , a log-scale posterior density will be used for sampling to avoid issues with arithmetic and machine rounding of small probabilities. Substituting Eqs. (17) and (16) into Eq. (11) and grouping terms yields the negative-log posterior,

$$-\ln[\mathbb{P}(\mathbf{x}, \boldsymbol{\theta}|\mathbf{b})] = \frac{\mathbf{e}^T \mathbf{e}}{2\sigma_e^2} + \sum_{i=1}^{d_x} \frac{(x_i - \mu_{x,i})^2}{2\sigma_{x,i}^2} + \sum_{i=1}^{d_\theta} \frac{(\theta_i - \mu_{\theta,i})^2}{2\sigma_{\theta,i}^2} + Z, \quad (20)$$

where Z is a constant that does not affect the topology of the distribution.

Markov chain Monte Carlo (MCMC) refers to a class of methods that can be used to sample a posterior PDF. As the number of samples becomes large, the density of samples becomes proportional to the target distribution. There are several MCMC algorithms available for sampling the posterior, including Gibbs, slice, and Metropolis–Hastings sampling. The Metropolis–Hastings algorithm generates steps with components along each dimension of \mathbf{x} and $\boldsymbol{\theta}$ at once, as opposed to Gibbs and slice sampling, which generate sequential samples along individual dimensions of the distribution. To allow for flexibility in the model input vectors, Metropolis–Hastings sampling was chosen. The Markov chain is initiated at the least-squares estimate of the QoI to minimize the “burn-in” time. For the LRS inference, in the case of the pressure, which was measured, its mean and variance are set by the data. The remaining prior widths are set by choosing physically sensible minimum and maximum ranges for the parameters. Then, the posteriors for the relevant parameters of the LRS inference are used to construct the priors for those variables in the LRS inference. The rest of the priors are constructed in the same way as above. MCMC chains in this work are run until the cumulative mean trace and correlation signals are determined to be approximately stationary.

It is often convenient to invoke an approximate form of the posterior, which bears the name of Laplace: Assuming that the posterior is smooth, strongly peaked, and unimodal, it may be modeled using a Gaussian PDF that is centered at the global maximum and fit to the local curvature at that point.³² The maximum is found by an optimization algorithm, and the Jacobian at that point is squared to approximate the Hessian and thereby estimate the inverse covariance matrix. The utility and validity of this model of $\mathbb{P}(\mathbf{b}|\mathbf{x}, \boldsymbol{\theta})$ for a sequential inference, such as LRS-calibrated LTS, is not known *a priori*. In addition, the effect of the accuracy of prior information on the ability to distinguish between candidate submodels is also unclear at this stage. These questions are explored in Secs. IV and V B.

D. Bayesian model selection

Up until now, our discussion of Bayesian inference has assumed a fixed model of the target process and thus a constant set of QoI. Many models are implicated in any estimation of the QoI, including the measurement model, physics models (such as the submodels in Sec. II), error distributions, priors, and hyperpriors, and we do not always know the best models to use, *a priori*. Bayesian model selection enables the comparison of models with incommensurate QoI. In this paper, we are particularly interested in the application of model selection to plasma submodels to deduce the presence of non-equilibrium behaviors and ensure robust estimation of the unknown parameters.

To start, it should be noted that any inference is implicitly conditional on the model. We make this dependence explicit and express an update using the i th model, M_i ,

$$\mathbb{P}(\mathbf{x}, \boldsymbol{\theta}|\mathbf{b}, M_i) = \frac{\mathbb{P}(\mathbf{b}|\mathbf{x}, \boldsymbol{\theta}, M_i)\mathbb{P}(\mathbf{x}, \boldsymbol{\theta}|M_i)}{\mathbb{P}(\mathbf{b}|M_i)}. \quad (21)$$

We can compute the so-called model-posterior PDF for M_i using Bayes' equation,

$$\mathbb{P}(M_i|\mathbf{b}) = \frac{\mathbb{P}(\mathbf{b}|M_i)\mathbb{P}(M_i)}{\mathbb{P}(\mathbf{b})}. \quad (22)$$

The “model likelihood,” $\mathbb{P}(\mathbf{b}|M_i)$, indicates the degree to which a model and one's prior assumptions are consistent with the measurement vector, \mathbf{b} . Therefore, the ratio of model likelihoods for candidate models M_i and M_j can be used to assess their relative plausibility,

$$B_{i,j} \equiv \frac{\mathbb{P}(M_i|\mathbf{b})}{\mathbb{P}(M_j|\mathbf{b})} = \frac{\mathbb{P}(\mathbf{b}|M_i)\mathbb{P}(M_i)}{\mathbb{P}(\mathbf{b}|M_j)\mathbb{P}(M_j)}, \quad (23)$$

where $B_{i,j}$ is the so-called Bayes factor. Bayes factors greater than one imply that the data in \mathbf{b} are more consistent with model i than model j , vice versa for $B_{i,j} < 1$; and one is indifferent to M_i and M_j when $B_{i,j}$ is equal to unity.

Assigning a prior distribution to a *model* requires comprehensive knowledge of the model space, which is rarely available. Absent specific information, we assume that all our submodels are equally

plausible: $\mathbb{P}(M_i) \approx \mathbb{P}(M_j)$ for all i and j . This assumption is used in conjunction with Eq. (21) to calculate the Bayes factor,

$$B_{i,j} \approx \frac{\mathbb{P}(\mathbf{b}|M_i)}{\mathbb{P}(\mathbf{b}|M_j)} = \frac{\iint \mathbb{P}(\mathbf{b}|\mathbf{x}, \boldsymbol{\theta}, M_i) \mathbb{P}(\mathbf{x}, \boldsymbol{\theta}, M_i) \, d\mathbf{x} \, d\boldsymbol{\theta}}{\iint \mathbb{P}(\mathbf{b}|\mathbf{x}, \boldsymbol{\theta}, M_j) \mathbb{P}(\mathbf{x}, \boldsymbol{\theta}, M_j) \, d\mathbf{x} \, d\boldsymbol{\theta}}. \quad (24)$$

Equation (24) may be approximated via the Monte Carlo method,

$$B_{i,j} \approx \frac{N_j}{N_i} \frac{\sum_{k=1}^{N_i} (2\pi\sigma_{e,i}^2)^{-d_b/2} \exp\left(-\frac{\mathbf{e}_i^T \mathbf{e}_i}{2\sigma_{e,i}^2}\right)}{\sum_{k=1}^{N_j} (2\pi\sigma_{e,j}^2)^{-d_b/2} \exp\left(-\frac{\mathbf{e}_j^T \mathbf{e}_j}{2\sigma_{e,j}^2}\right)}. \quad (25)$$

In this expression, N_i and N_j are the number of samples drawn from $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta}|M_i)$ and $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta}|M_j)$, respectively, which are used to evaluate the numerator and denominator, in turn. The error variances, $\sigma_{e,i}^2$ and $\sigma_{e,j}^2$, may be model-specific, in principle, but we use a single variance for all our plasma submodels.

We take the logarithm of Eq. (25) to provide numerical stability. Our results are presented in terms of $\log(B_{i,j})$ such that positive values indicate stronger support for model M_i than M_j , negative values suggest the opposite, and zero indicates indifference.

The Monte Carlo estimation of the log-evidence is effective when the prior PDF plays a large role in an inference. However, when the prior is very diffuse compared to the likelihood, this technique leads to an excessive number of samples in low-density regions of the model evidence, which poses a large computational cost. Efficient estimation of the log-evidence is an active area of research; see Refs. 34–36 for alternative methods to estimate this quantity.

IV. DISTINGUISHING VELOCITY DISTRIBUTION FUNCTIONS WITH BAYESIAN MODEL SELECTION

A. Synthetic data, priors, and test matrix

As stated above, we put independent Gaussian distributions on all our QoI and nuisance parameters. In the general case, a combination of priors, and specifically a log-uniform prior for the electron number density, may be preferred, as was recommended in Ref. 17. In addition, it is assumed that an LRS Bayesian inference or least-squares inversion has been used to provide the mean and width of the priors for η and λ_i , which are nuisance variables in the LTS inversion. Table IV outlines the $3\sigma_{x,i}$ ranges for the prior PDFs. These values were pulled, stretched, and rounded from a previous least-squares signal inversion of collected LTS data, as representative of an acceptable approximation to the posterior or “true” prior.

In order to explore the effects of SNR and prior uncertainty on the “distinguishability” of candidate submodels, synthetic data are generated at several SNRs and degrees of prior “misinformation.” We define the SNR as the ratio of the generated data’s signal power to the noise power,

$$\text{SNR} \equiv \frac{E[\max(\hat{\mathbf{b}})^2]}{E[d_b^{-1}(\hat{\mathbf{b}} - \tilde{\mathbf{b}})^2]}, \quad (26)$$

where $E(\cdot)$ is the expectation operator. Using this formula, we generate synthetic data with a prescribed SNR by setting the variance to

$$\sigma_{\text{noise}}^2 = \frac{\text{mean}[\max(\mathbf{b})^2]}{\text{SNR}}. \quad (27)$$

For ground truth values of \mathbf{x} and $\boldsymbol{\theta}$, the simulated measurements are given by

$$\tilde{\mathbf{b}} = \hat{\mathbf{b}}(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{e}_{\text{synth}}, \quad (28)$$

where $\mathbf{e}_{\text{synth}}$ is an independent and identically distributed random variable with zero mean and variance σ_{noise}^2 .

We characterize the performance of Bayesian model selection as a function of the accuracy and precision of the prior. To do this, the prior PDFs can be widened and offset as a function of a percentage of the original range and mean,

$$\mu_{x,i}^{\text{shift}} = \mu_{x,i} \left(1 + \frac{\kappa_i^{\text{shift}}}{100}\right) \quad (29)$$

and

$$\sigma_{x,i}^{\text{stretch}} = \mu_{x,i}^{\text{shift}} \pm \frac{\text{range}(x_i)}{2} \left(1 + \frac{\kappa_i^{\text{stretch}}}{100}\right), \quad (30)$$

where κ_i^{shift} and $\kappa_i^{\text{stretch}} \in [-100, 100]$. These parameters allow for independent stretching and shifting of the prior PDF within a $\pm 3\sigma_{x,i}$ range with respect to the original, data-generating prior’s $\pm 3\sigma_{x,i}$ range.

The cases presented in Secs. IV B–IV D were generated in the following way. First, for a given combination of $\kappa_{x,i}^{\text{shift}}$ and $\kappa_{x,i}^{\text{stretch}}$, baseline values of \mathbf{x} and $\boldsymbol{\theta}$ were drawn from $\mathbb{P}(\mathbf{x}, \boldsymbol{\theta})$ according to Table IV. These were used to generate a synthetic signal, $\hat{\mathbf{b}}(\mathbf{x}, \boldsymbol{\theta})$, using M_1 and M_2 from Table III. Models M_1 and M_2 were chosen for this comparison because they are parameterized with the same parameter vector \mathbf{x}^T . Then, at each of the 1150 SNR values between 1 and 100, the data were perturbed according to Eq. (28). The value of $\ln(B_{2,1})$ was calculated at each SNR two times, each using 500 samples from the perturbed prior. Similarly, to explore the ability to detect deviations from the submodels, model M_3 (a binary Maxwellian–Druryvesteyn plasma) was used to generate the data at a given value of γ_{MD} , according to the process outlined above. In these

TABLE IV. Description of the $3\sigma_{x,i}$ ranges for the submodel parameters used in the synthetic LTS prior distributions.

Parameter	Minimum	Maximum
γ_{MD}	0.25	0.75
T_e	2	3
T_e^{M}	2	3
T_e^{D}	1	2
n_e	1×10^{18}	5×10^{18}
v_d	−50	50
η	0.160	0.190
λ_i	531.9	532.1

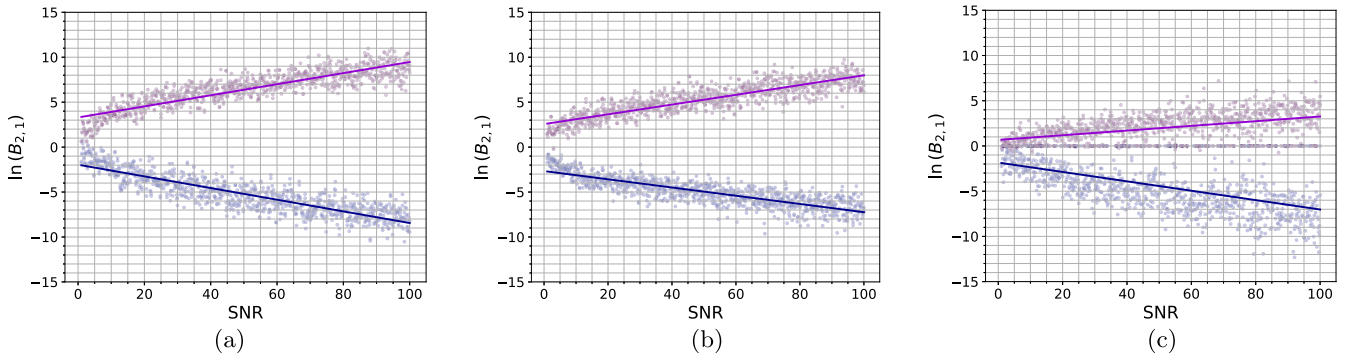


FIG. 2. $\ln(B_{2,1})$ vs SNR over priors with increasing misinformation: (a) $\kappa_{x,i}^{\text{stretch}} = 0$, (b) $\kappa_{x,i}^{\text{stretch}} = 100$, and (c) $\kappa_{x,i}^{\text{stretch}} = 300$. Synthetic data generated from model M1 are in blue, and synthetic data generated from model M2 are in purple. The dots indicate the calculated value of $\ln(B_{i,j})$, and the lines indicate a line of best fit. $|\ln(B_{i,j})|$ substantially greater than zero indicates that the Bayesian framework is able to correctly identify the right model.

cases, 25 values of γ_{MD} between 0 and 1 were used, with 25 values of SNR between 1 and 100.

B. Effects of prior uncertainty

To understand the effect of uncertainty in the prior (i.e., based on its width) on one's ability to distinguish candidate submodels, $\ln(B_{2,1})$ is calculated using priors at varying levels of width misinformation, with $\kappa_{x,i}^{\text{stretch}} \in \{0, 100, 300\}$ for all x_i .

The results are presented in Fig. 2. We expect data generated with M_1 to produce $\ln(B_{2,1}) < 0$ and data generated with M_2 to produce $\ln(B_{2,1}) > 0$. In addition, we expect the magnitude of $\ln(B_{2,1})$ to decrease with decreasing SNR. At each SNR, a line of best fit was computed by averaging the two values of $\ln(B_{2,1})$ at each SNR and then fitting a line. This is shown in Fig. 2. In Figs. 2, 4, and 6, all purple data points and the best fit line being positive and all blue points and the best fit line being negative indicate that, in all cases, the correct model was inferred.

In the base case (a), in which the generating distribution functions are used as the priors (i.e., there is no prior misinformation), the data demonstrate that the proposed method can distinguish

between candidate submodels over the entire range of SNRs from 1 to 100. Furthermore, there is strong support for the correct model when the $\text{SNR} \geq 5$. In cases (b) and (c), where there is prior misinformation in the form of a stretched PDF, the ability to distinguish between the correct model is decreased compared to (a); case (c) occasionally returns indistinguishable values of $\ln(B_{2,1})$, which were automatically set to 0 in the data processing code.

Figure 3 explores the ability to detect deviations from each submodel using the binary model M_3 . Similar to the pure-Maxwellian and pure-Druyvesteyn cases, support for the model drops with decreasing SNR. However, the ability to distinguish between the “degree of Maxwellianization” is maintained over the SNR range considered. This indicates that the Bayesian model selection framework is able to correctly distinguish between the candidate submodels and can distinguish between interpolated submodel candidates accurately in accordance with the interpolation parameter γ_{MD} over decreasingly informative priors. However, the slope of $\ln(B_{2,1})$ vs γ_{MD} decreases in magnitude with increasing prior widths [cases (b) and (c)], indicating that the models are harder to distinguish with wider priors.

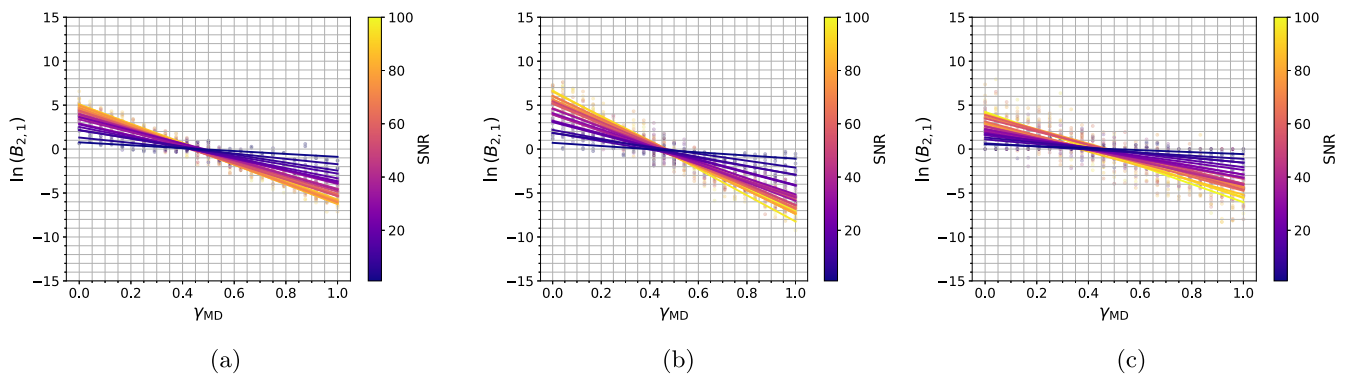


FIG. 3. $\ln(B_{2,1})$ vs γ_{MD} at varying SNR levels over decreasingly certain priors for model M_3 : (a) $\kappa_{x,i}^{\text{stretch}} = 0$, (b) $\kappa_{x,i}^{\text{stretch}} = 100$, and (c) $\kappa_{x,i}^{\text{stretch}} = 300$. The dots indicate the calculated value of $\ln(B_{i,j})$, and the lines indicate a line of best fit. γ_{MD} characterizes how “Maxwellian” the plasma is.

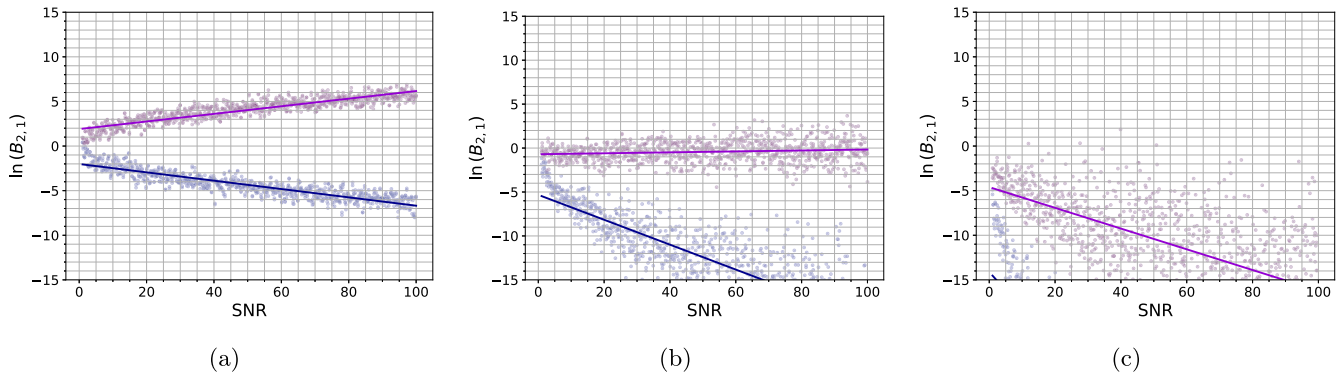


FIG. 4. $\ln(B_{2,1})$ vs SNR for decreasingly accurate priors: (a) $\kappa_{x_i}^{\text{shift}} = 1$, (b) $\kappa_{x_i}^{\text{shift}} = 20$, and (c) $\kappa_{x_i}^{\text{shift}} = 30$. Synthetic data generated from model M1 are in blue, and synthetic data generated from model M2 are in purple. The dots indicate the calculated value of $\ln(B_{i,j})$, and the lines indicate a line of best fit. $|\ln(B_{i,j})|$ substantially greater than zero indicates that the Bayesian framework is able to correctly identify the right model.

C. Effects of prior inaccuracy

We wish to understand the effect of the accuracy of the mean of a prior PDF on the ability to distinguish candidate submodels as a function of SNR. To do this, $\ln(B_{2,1})$ is calculated using priors at varying levels of prior mean misinformation, with $\kappa_{x_i}^{\text{shift}} \in \{1, 20, 30\}$ for all x_i except λ_i . The incident wavelength was fixed during this test, as the signal inversion is extremely sensitive to the center wavelength. In addition, any shift past a percentage of 30 on any of the other variables led to uncomputable submodels, with the data processing code returning $\ln(B_{2,1}) = 0$ overall SNR values. The results are presented in Fig. 4.

As opposed to the uniform stretching of the prior information, Fig. 4 indicates that shifting the prior mean by as little as 1% leads to a rapid drop in support for the correct model, which is shown by the spread of the raw values of $\ln(B_{2,1})$. This is true even at moderately high SNRs, above around 50, and gets progressively worse at larger shifts in the prior means from the true prior. Even small shifts in the prior mean with respect to the true value

lead to persistent effects on the residuals that skew $\ln(B_{2,1})$, leading to the stark contrast between Figs. 4 and 2. This sensitivity to prior inaccuracy can be indicative of a very strong prior distribution compared to the information in the data. In such cases, it is often desirable to collect higher SNR data and use wider prior distributions, especially when considering extensions to higher-order mixtures of distribution functions.

Accurate discernment of the degree of Maxwellianization is evident at small prior information shifts from Fig. 5. However, again, at large shifts from the true mean, the ability to discern the degree of Maxwellianization (γ_{MD}) is lost, even at high SNRs.

The data above assume perfect information about the laser intensity, λ_i . However, the same trends are also observed when the mean of λ_i is subjected to shifts of $\kappa_{\lambda_i}^{\text{shift}} \in \{0.01, .5\}$, indicating the sensitivity of the model selection on a very accurate estimate of the center wavelength. This means that when the laser wavelength is included as an unknown parameter, even sub-percentage errors in the estimates of the center wavelength lead to the complete loss of model distinguishability.

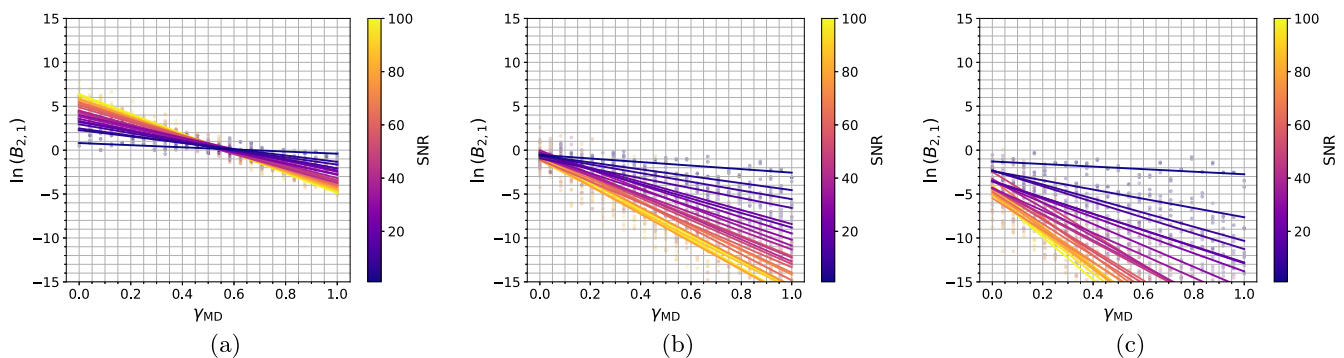


FIG. 5. $\ln(B_{2,1})$ vs γ_{MD} at varying SNR levels over decreasingly informative priors with model M_3 : (a) $\kappa_{x_i}^{\text{shift}} = 1$, (b) $\kappa_{x_i}^{\text{shift}} = 20$, and (c) $\kappa_{x_i}^{\text{shift}} = 30$. The dots indicate the calculated value of $\ln(B_{i,j})$, and the lines indicate a line of best fit. γ_{MD} characterizes how “Maxwellian” the plasma is.

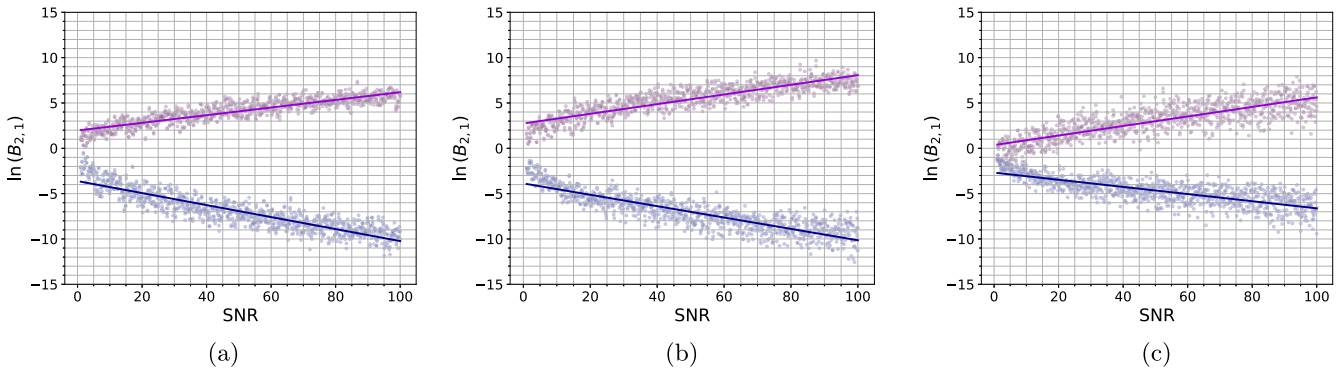


FIG. 6. $\ln(B_{2,1})$ vs SNR over decreasingly informative priors: (a) $(\kappa_{x,i}^{\text{stretch}}, \kappa_{x,i}^{\text{shift}}) = (10, 2)$, (b) $(\kappa_{x,i}^{\text{stretch}}, \kappa_{x,i}^{\text{shift}}) = (25, 5)$, and (c) $(\kappa_{x,i}^{\text{stretch}}, \kappa_{x,i}^{\text{shift}}) = (100, 20)$. Synthetic data generated from model M1 are in blue, and synthetic data generated from model M2 are in purple. The dots indicate the calculated value of $\ln(B_{i,j})$, and the lines indicate a line of best fit. $|\ln(B_{i,j})|$ substantially greater than zero indicates that the Bayesian framework is able to correctly identify the right model.

D. Interactions of inaccuracy and uncertainty

To understand the effect of simultaneously inaccurate and uncertain prior PDFs on the ability to distinguish between candidate submodels, $\ln(B_{2,1})$ is calculated using priors at varying levels of prior mean and width misinformation, with $\kappa_{x,i}^{\text{stretch}} \in \{10, 25, 100\}$ and $\kappa_{x,i}^{\text{shift}} \in \{2, 5, 20\}$ for all x_i except for λ_i . λ_i was fixed during this test for the reason presented above. The results are shown in Fig. 6.

Overall, the ability to accurately distinguish between candidate submodels is preserved over SNRs greater than 5, getting worse as the priors diverge from the true priors. However, there is an artificially inflated and artificially deflated support for the Maxwellian and Druyvesteyn models, respectively, as the priors diverge from the true priors, indicating that poorer initial construction of the information in the priors results in an incorrect preference for a Maxwellian submodel. These trends are mirrored in Fig. 7, indicating that the ability to discern small deviations from Maxwellian or Druyvesteyn submodels is lost even at values of $\gamma_{\text{MD}} = 0$, with an inflated preference for the Maxwellian submodel. In short, even for small deviations from a Druyvesteyn submodel, incorrect priors may artificially

prefer a Maxwellian submodel, even if they are wide enough to include the true mean.

The cases presented in Secs. IV C, IV B, and IV D indicate the necessity of well-constructed priors. However, the uncertainty bounds and correlations provided by a least-squares inference need to be validated by comparing the posterior generated using the Laplace approximation to the posterior PDF using MCMC sampling. This is explored in Sec. V B. In addition, the artificial preference of a Maxwellian submodel, even when the candidate submodels being compared are parameterized with the same \mathbf{x} , brings into question how the fidelity of the model, or models with increased parameter spaces, affects the selection of the model. This question is explored in Sec. IV E.

E. Transdimensional model selection

To understand the effect of model fidelity on the ability to distinguish between candidate submodels as a function of SNR, data are generated using models M_1, M_2, M_3 , and M_4 according to the

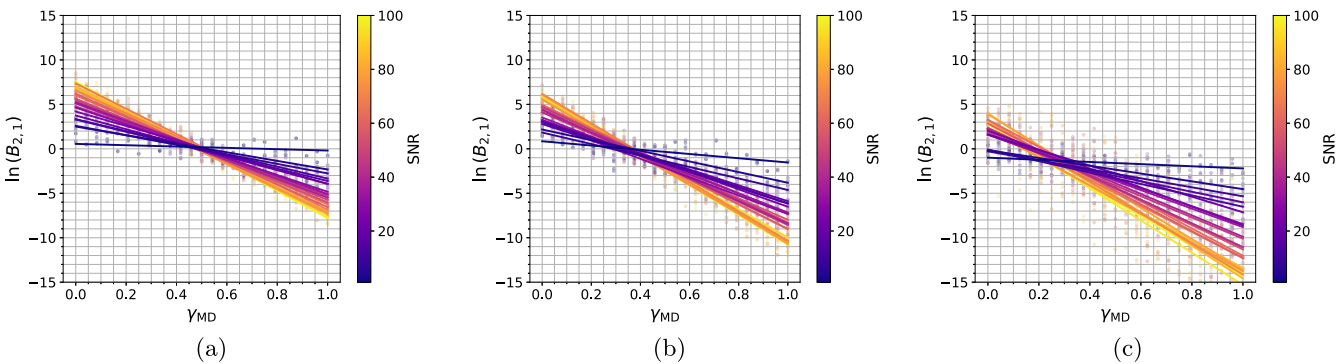


FIG. 7. $\ln(B_{2,1})$ vs γ_{MD} at varying SNR levels over decreasingly informative priors: (a) $(\kappa_{x,i}^{\text{stretch}}, \kappa_{x,i}^{\text{shift}}) = (10, 2)$, (b) $(\kappa_{x,i}^{\text{stretch}}, \kappa_{x,i}^{\text{shift}}) = (25, 5)$, and (c) $(\kappa_{x,i}^{\text{stretch}}, \kappa_{x,i}^{\text{shift}}) = (100, 20)$. The dots indicate the calculated value of $\ln(B_{i,j})$, and the lines indicate a line of best fit. γ_{MD} characterizes how “Maxwellian” the plasma is.

parameters in Table IV. Then, two values of $\ln(B_{2,1})$, $\ln(B_{3,1})$, and $\ln(B_{4,1})$ are calculated using 1000 samples drawn from the actual data generating priors (true priors) at 1000 SNRs between 1 and 100.

The results are shown in Fig. 8. In case (a), with data generated via M_1 , the $\ln(B_{i,j})$ indicate that, in the case where the true data are Maxwellian, regardless of which model is being compared to the Maxwellian model, the model selection framework is able to discern the correct model.

In case (b), with data generated with M_2 , $\ln(B_{2,1}) > \ln(B_{3,1}) > \ln(B_{4,1})$, with the spread of $\ln(B_{2,1})$ and $\ln(B_{3,1})$ both being less than $\ln(B_{4,1})$. This indicates that the model selection framework was able to correctly distinguish between the models. This is due to the fact that $\ln(B_{3,1})$ was a single temperature, multi-model distribution. The values of γ_{MD} near zero would effectively create the same distribution as model M_2 , leading to a higher preference than model M_4 because the temperatures were not allowed to overlap in the data-generating priors.

In case (c), with data generated with M_3 , $\ln(B_{3,1}) > \ln(B_{2,1}) > \ln(B_{4,1})$ also indicates the ability to correctly determine the model, as well as the preference for M_1 over the higher parameter space M_4 . This is likely due to the non-overlapping temperatures for M_4 , making it impossible to generate an equivalent single temperature distribution.

Finally, in case (d), with data generated with M_4 , $\ln(B_{4,1}) > \ln(B_{3,1}) > \ln(B_{2,1})$. Again, the ability to discern the correct model is clear. The order of preference is also expected. That is, because M_1 and M_2 are single temperature and density fraction models, they are not able to properly fit the intermediate shapes created by the superposition of EVDFs with multiple temperatures and number density fractions.

Cases (a)–(d) indicate that the assumption of equally probable priors on the models, $\mathbb{P}(M_i) \approx \mathbb{P}(M_j)$, used in deriving Eq. (25) is valid over the fidelity of the models evaluated, as compared to the Maxwellian submodel. In addition, increasing model fidelity does result in a preference for the correct submodel that best describes the data, validating the model selection framework.

V. APPLICATION OF BAYESIAN LTS TO NS-PULSED PLASMAS

We now demonstrate the Bayesian LTS model selection and signal inversion process using experimental data. An LRS-calibrated LTS system was developed for an optically accessible, stainless-steel, pressure-controlled vacuum cell that was designed to generate pulsed and steady-state DC discharge plasmas. However, at high steady-state bias voltages, arcing events prevent the collection of spectra due to plasma instability; the ns-pulsed plasma was chosen for study here.

A. Experimental setup

The plasma was generated by kV-level biases between a set of stainless steel electrodes with 24 mm diameter, 26.53 mm spacing. The electrodes were floating with respect to the stainless steel body of the vacuum cell, which was tied to the electrical ground. Power feedthroughs isolated the discharge electrodes from the vacuum test cell, ensuring that their base configuration was floating with

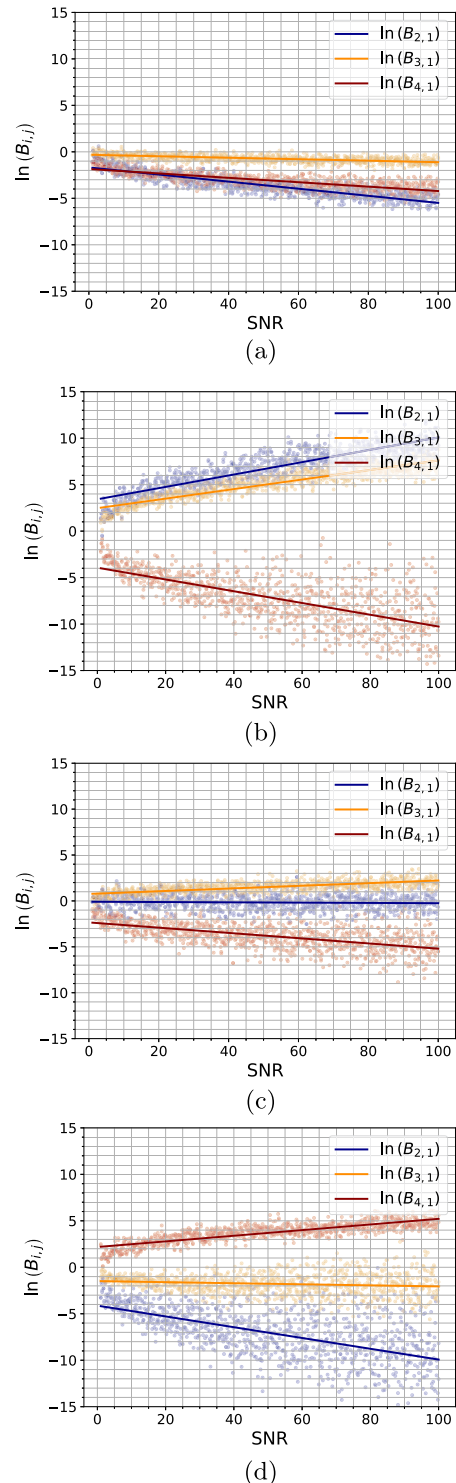


FIG. 8. $\ln(B_{i,j})$ vs SNR over several data generation cases: (a) data generated with M_1 , (b) data generated with M_2 , (c) data generated with M_3 , and (d) data generated with M_4 . $\ln(B_{2,1})$ (blue), $\ln(B_{3,1})$ (orange), $\ln(B_{4,1})$ (red). The dots indicate the calculated value of $\ln(B_{i,j})$, and the lines indicate a line of best fit.

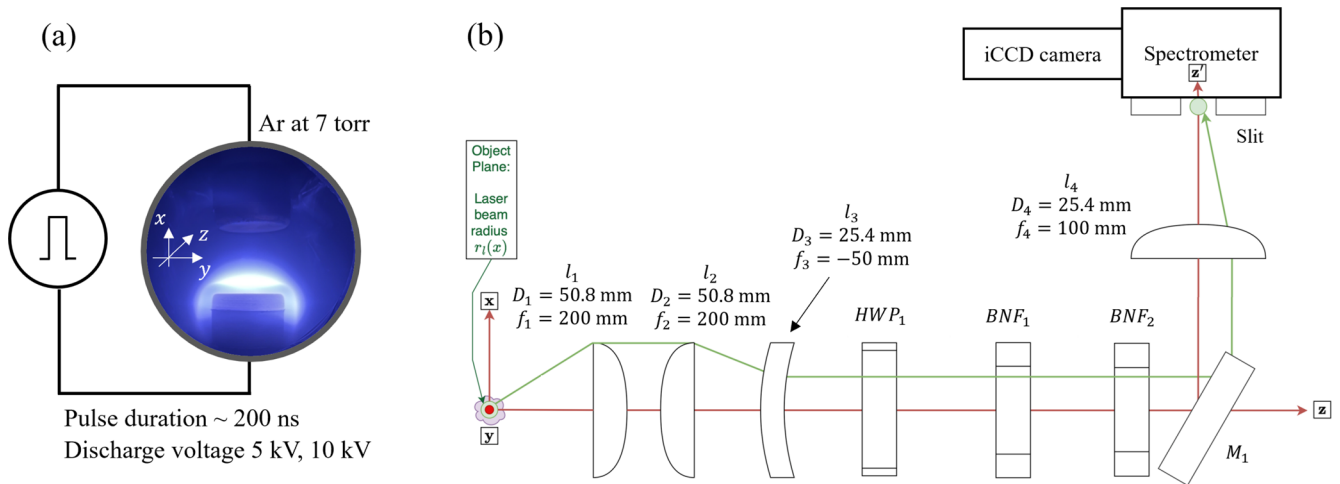


FIG. 9. (a) A nanosecond pulsed discharge setup and a sample discharge image. The interrogation laser beam propagates in the x direction and is polarized in the y direction, and the scattering is collected in the z direction. (b) A diagram of the collection system optical setup. M , l , HWP , and BNF are acronyms for mirrors, lenses, half waveplates, and Bragg notch filters, respectively.

respect to the vacuum cell and, therefore, with respect to the electrical ground. The voltage pulse width was 200 ns, with a 100 Hz frequency. A sample discharge image and a setup schematic are shown in Fig. 9(a).

Thomson and Raman scattering were stimulated by the second harmonic output of a pulsed Nd:YAG laser (Quantel Q-smart 850, 200 mJ/pulse). The interrogation was steered, shaped, polarization-aligned, and focused into the test section using a series of mirrors, lenses, half waveplates, and polarizing beam splitter cubes and then transmitted to the plasma through Brewster windows on the vacuum cell. The scattered light was collected, relayed, and imaged onto the spectrometer using achromatic optics. A diagram of the detection system optical setup is presented in Fig. 9(b). The system detection solid angle, $\Delta\Omega$, was optimized by maximizing the collection angle of the first collection lens while using the two volume Bragg-grating notch filters (BNFs) as marginal- and chief-ray-limiting apertures. Two BNFs (OptiGrate BNF-532-OD4) were used to filter the center-line laser Rayleigh scattering and reflections to maximize the useable length along the interrogation beam waist; see Ref. 19 for details. LRS and LTS signals were detected using an intensified camera (Princeton Instruments PM4-1024i-HB-FG-18-P46, 15-ns gate, 3000 shots) coupled to a spectrometer (Princeton Instruments Isoplane 320A, 2400 l/mm grating). A delay generator (Berkeley Nucleonics BNC-577) controlled the relative timing of the detector gate delay with respect to the laser Q-switch to maximize the signal collected on the detector. The delay generator also controlled the relative timing of the laser and ns-pulsed power supply trigger, thus allowing for measurements at different times in the pulsed plasma cycle. Descriptions of previous iterations of the plasma and LTS facilities at Georgia Tech and Texas A&M University can be found in Bak *et al.*¹⁹ A diagram of the detection system optical setup is presented in Fig. 9. Please see Ref. 37 for a detailed description of the LTS and plasma cell used in this work.

LRS was collected at N_2 neutral background pressures of 10, 7, and 4 Torr. Then, a background spectrum, B_r , was taken at

the 0.8 Torr base pressure of the system. The subtraction of these two signals left the pure Raman spectrum to be analyzed. For the LTS measurements, plasma was generated at 7 Torr Ar neutral background pressure with two discharge voltages, 5 and 10 kV. Under each LTS condition, measurements were obtained with the plasma on, laser on; plasma off, laser off; and plasma off, laser on. These signals were appropriately subtracted, following Ref. 10, to yield the LTS spectrum to be analyzed.

B. Bayesian inference and model selection

In order to compare the Laplace approximation described in Sec. III C to that of a full MCMC-sampled Bayesian inference, the nonlinear least-squares method was used to determine the relevant QoI and their variance. The parameter vectors used were

$$\mathbf{x}_{LS}^R = [T_g, \tau, \eta, \lambda_i] \quad (31)$$

and

$$\mathbf{x}_{LS}^T = [T_e, n_e, v_d] \quad (32)$$

for the least-squares optimization. Variables that have the same effect on the inference of interest—such as p_g and η on the Raman inference, or η and n_e , or v_d and λ_i on the Thomson inference—are not easily decoupled, and their uncertainty is not included using the standard least-squares residual methods. At each target Raman pressure, the expected value and the variance of p_g were estimated using the mean and square of the standard deviation of the pressure trace. Then, a scaled and shifted Gaussian was fit to the B_r signal to provide initial estimates of τ and λ_i . The initial estimate of the gas temperature, T_g , was chosen to be 295.15 K. The temperature bounds were chosen to be $\pm 10\%$, and the other variables in the LRS LS vector were allowed to range $\pm 100\%$.

The least-squares method was used to determine \mathbf{x}_{LS}^R and its variance, $\delta\mathbf{x}_{LS}^R$, in accordance with a 95% confidence interval procedure,³⁸ after approximating the least-squares covariance matrix. The least-squares covariance matrix, Γ_{LS} , is approximated using the Jacobian, \mathbf{J} , from the least-squares algorithm in the Python package,

$$\Gamma_{LS} \approx \frac{(\mathbf{J}^T \mathbf{J})^{-1} \text{diag}(\mathbf{e})^2}{N_\lambda - \text{length}(\mathbf{x}_{LS})}, \quad (33)$$

where $\mathbf{A}^2 = \mathbf{A}\mathbf{A}$. The same method is used to quantify the uncertainty in the LTS case. Two fitted spectra are presented in Fig. 10.

In order to calculate \mathbf{x}_{LS}^T , λ_i and η are assumed to be constant at their optimal value for reasons stated previously, and their uncertainty is not propagated forward. This leads to a possible solution for

each individually collected Raman spectrum, although in the ideal case, they should all produce the same value of λ_i and η . There can be a small to possibly large variation in these parameters, specifically η , which can produce very different results in the inversion of the Thomson signal for n_e and v_d . Ignoring the uncertainty in η and comparing it to Bayesian inference will provide insights into when it is necessary to propagate the uncertainty in η forward and when it is appropriate to simply use the uncertainty from the least-squares residuals.

The value of the relevant LS-LRS inference parameters η and λ_i , where each \mathbf{x}_{LS}^R is evaluated, is shown in Fig. 11. The source of the 10% variation in the mean values of η in Fig. 11 can be attributed to several factors. The vacuum cell pump down compression is expected to be of μm order. The laser itself has an energy stability of 4% and pointing stability of at most μrad , of which the combination could possibly lead to micro-movements of the beam height and lateral location in the test section compounding into the observed 10% variation in η .

Generally, uncertainty in λ_i is due to relative misalignment of the image of the laser beam on the detector relative to the calibrated

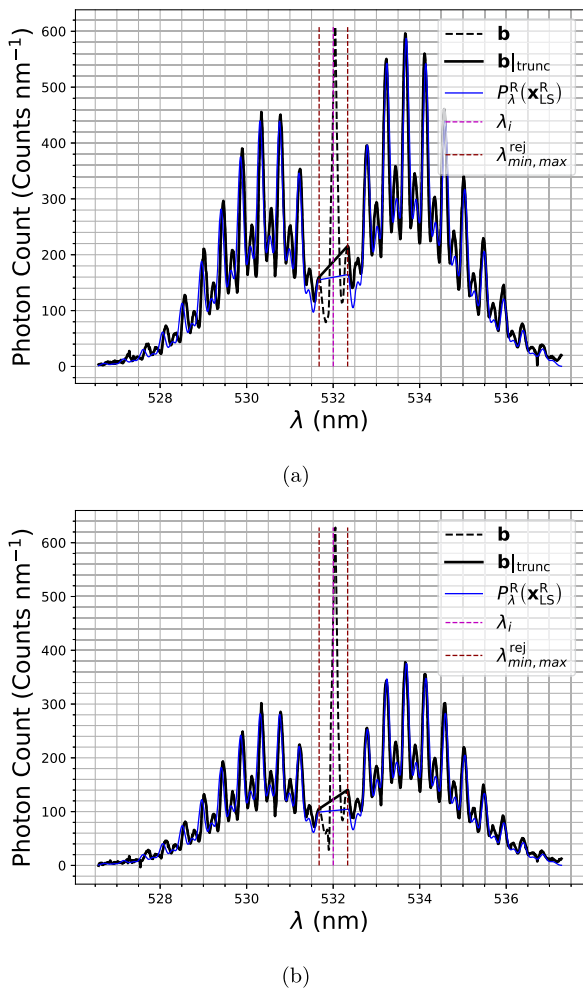


FIG. 10. Raman least-squares fit at two background pressures: (a) $p_g = 10$ Torr and (b) $p_g = 7$ Torr. The data and rejected data are presented in black solid and black dashed lines, respectively. The estimated center wavelength is presented with a purple dashed line, and the rejected data region is presented with red dashed lines.

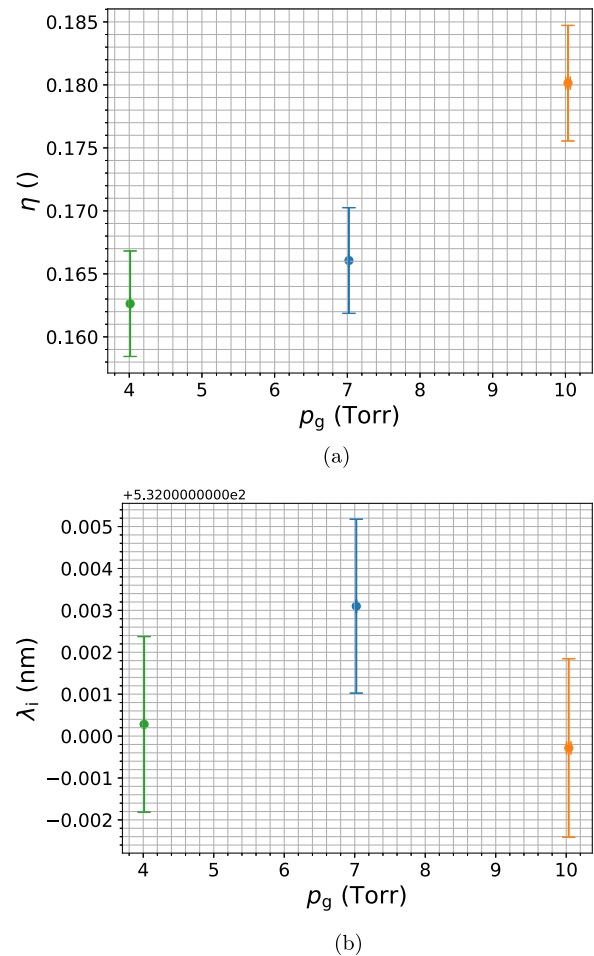


FIG. 11. Least-squares inferred parameters vs p_g : (a) η and (b) λ_i .

center wavelength pixel on the detector. This is why λ_i is included as a QoI. When the system has tight alignment, random fluctuations in the calculated values between Raman signals are expected, but like in our case, the fluctuations will be much less than 1%. In addition, micro-movements of the beam perpendicular to the spectrometer slit would introduce small changes in the calculated λ_i from the LRS inversion. This can be eliminated by assuming that λ_i is 532 nm. That assumption may allow for better relative bulk velocity measurements but will likely result in less accurate FWHM, neutral gas temperature, and calibration constant measurements when misalignment is prevalent.

At each value of η and λ_i from the LS-LRS, the bounds used for the LS-LTS inversion were $T_{e_0} = 10$ eV, $n_{e_0} = 1 \times 10^{20} \text{ m}^{-3}$, and

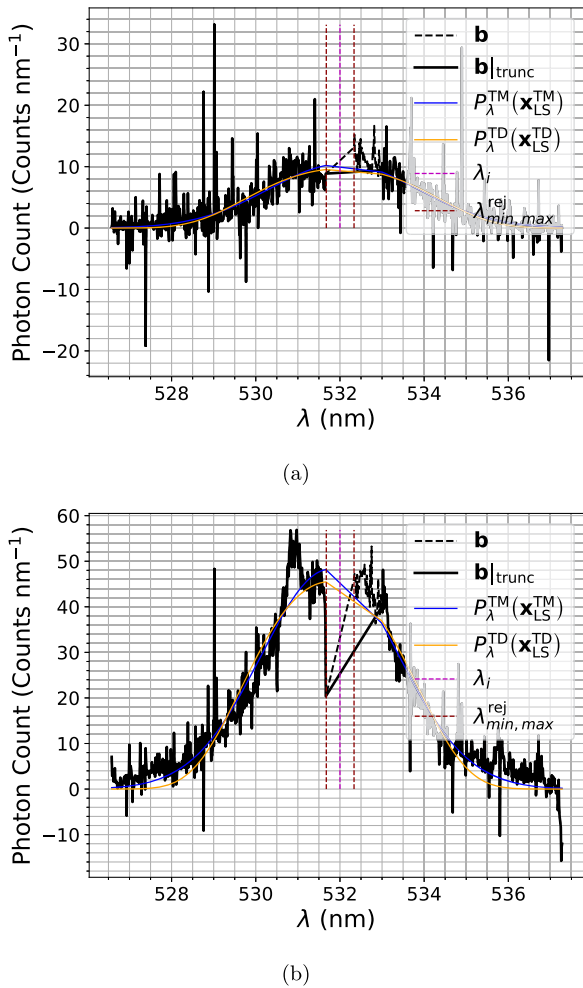


FIG. 12. Thomson least-squares fit at two discharge voltages: (a) $V_D = 5$ kV and (b) $V_D = 10$ kV. The data and rejected data are presented in black solid and black dashed lines, respectively. The estimated center wavelength is presented with a purple dashed line, and the rejected data region is presented with red dashed lines. The SNR estimates using the Maxwellian and Druyvesteyn submodels are 4 and 5 in (a) and 15 and 14 in (b), respectively.

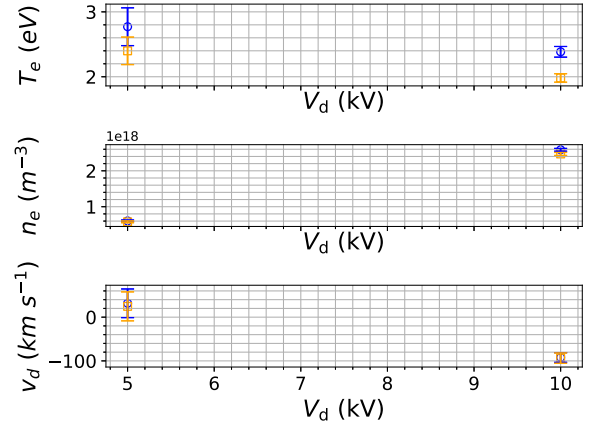


FIG. 13. Individual elements of the Thomson QoI \mathbf{x}_{LS}^T vs V_D . The Maxwellian submodel is presented in blue, and the Druyvesteyn submodel is presented in orange.

$v_d = 0 \text{ km s}^{-1}$, all bounded by $\pm 100\%$. The collected Thomson spectra are presented in Fig. 12. The QoI calculated from each of the 7 Torr Raman conditions is presented in Fig. 13.

The uncertainty in the calculated LRS parameters that are nuisance parameters in the LS-LTS inference at a given pressure has no effect on the electron parameters because there was no way of incorporating that information using the LS residual method directly without additional uncertainty propagation. However, the predicted electron number density shows a large dependence on the predicted optical efficiency constant η across pressures under each given LS-LRS condition. In addition, the predicted value and uncertainty are highly dependent on the model chosen to invert the signal of interest. This is likely due to the inability of the LS residual method to robustly include information such as the LRS uncertainty profile into the LTS analysis. This further motivates our more robust Bayesian method for including the relevant uncertainty profile from the LRS inference into that Thomson signal inversion.

Hence, the Bayesian inference framework was implemented for the LTS submodels M_1 and M_2 . The general methodology follows that outlined in Sec. III. The Bayesian inference model parameter vectors are taken as the same vectors presented in Sec. II. The priors are described in Table V, and the results are presented below. These ranges set the $3\sigma_{x,i}$ values for the Gaussian prior PDFs described in Eqs. (17)–(19).

A one million-element long Markov chain was constructed using a custom Metropolis–Hastings algorithm implemented in

TABLE V. Description of the $3\sigma_{x,i}$ ranges for the LRS case at 7 Torr.

Parameter	Minimum	Maximum
p_g	$0.99 * p_{g_0}$	$1.01 * p_{g_0}$
T_g	250	350
τ	0	0.5
η	0	1
λ_i	531	533

Python. This system was used to sample the log posterior for the 7 Torr Raman scattering spectrum. The chain was started at the LS-MAP estimate to reduce the burn-in time. The marginal and joint posterior distribution functions are presented in Fig. 14.

The prior was sampled as a set of independent Gaussian distributions, as shown in Fig. 14 in orange. The least-squares posterior was constructed as follows: given that the pressure, p_g , was not included in the least-squares inference, the prior was sampled. For the rest of the variables included in \mathbf{x}_{LS}^R , the least-squares MAP estimate and the covariance matrix were used to sample from a multivariate Gaussian distribution; this is shown in Fig. 14 in red. The Markov chain samples are shown in steel blue.

Across all cases except for the neutral gas pressure, the width of the priors is larger than that of the posterior estimates. In addition, the priors used for Bayesian inference covered a wider range than those for the least-squares case. However, least-squares posterior estimates for the LRS case are consistent with the Bayesian inference estimate. The case presented here covers the most conservative estimate of the priors, i.e., the case where the priors were the widest. This shows that even for generally uninformed priors, the least-squares provides a comparable uncertainty and parameter correlations.

Similarly, in the Thomson case, for both models M_1 and M_2 , a one million element length Markov chain was sampled. The priors for the nuisance parameters of the LS inference were set from

the posterior of the LRS inference. The priors for the LS QoI are described in Table VI, and the results are presented below.

Figure 15 shows the marginal and joint posterior PDFs for both the Druyvesteyn and Maxwellian models, with the Maxwellian model in blue and the Druyvesteyn model in orange, their LS posterior estimates in purple and red, and their prior estimates in green. Again, the MAP estimates and uncertainty bounds seem to be in agreement, showing the utility and necessity as shown in Sec. IV for accurate prior information. The joint posteriors show a divergence in the correlations between variables. This is due to the addition of the nuisance parameter uncertainties in the probability model that were not available with simple least-squares inversion. This indicates that the propagation of uncertainty, even for staged

TABLE VI. Description of the $3\sigma_{x,j}$ ranges for the LTS case at 7 Torr neutral background pressure 10 kV.

Parameter	Minimum	Maximum
T_e^M	0	100
T_e^D	0	100
n_e	1×10^{16}	5×10^{20}
v_d	-1000	1000

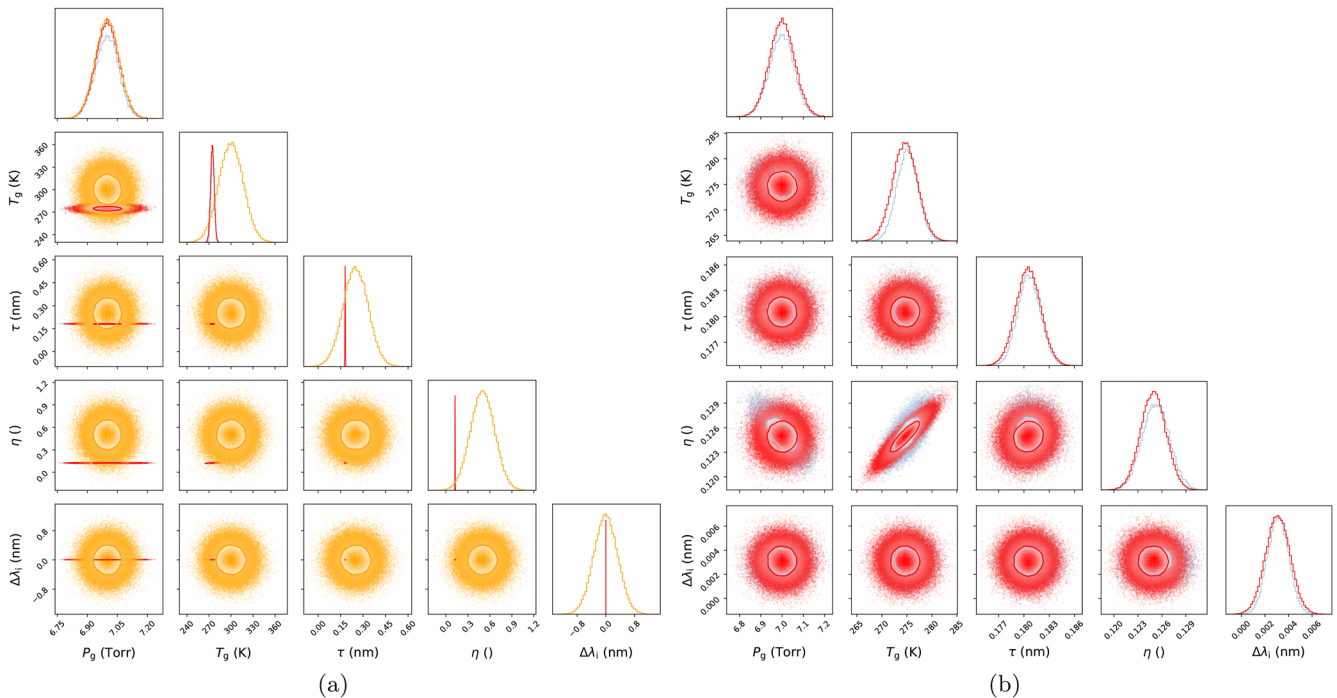


FIG. 14. LRS marginal and joint posterior distribution functions for all LRS model inputs along with the prior and least-squares posterior estimates. (a) Samples from the prior are in orange, and the Bayesian and LS posteriors are in steel blue and red, respectively. (b) The same without the prior. For the sake of clarity, each joint posterior displays a single Mahalanobis contour containing the most probable 11.8% of the distribution.

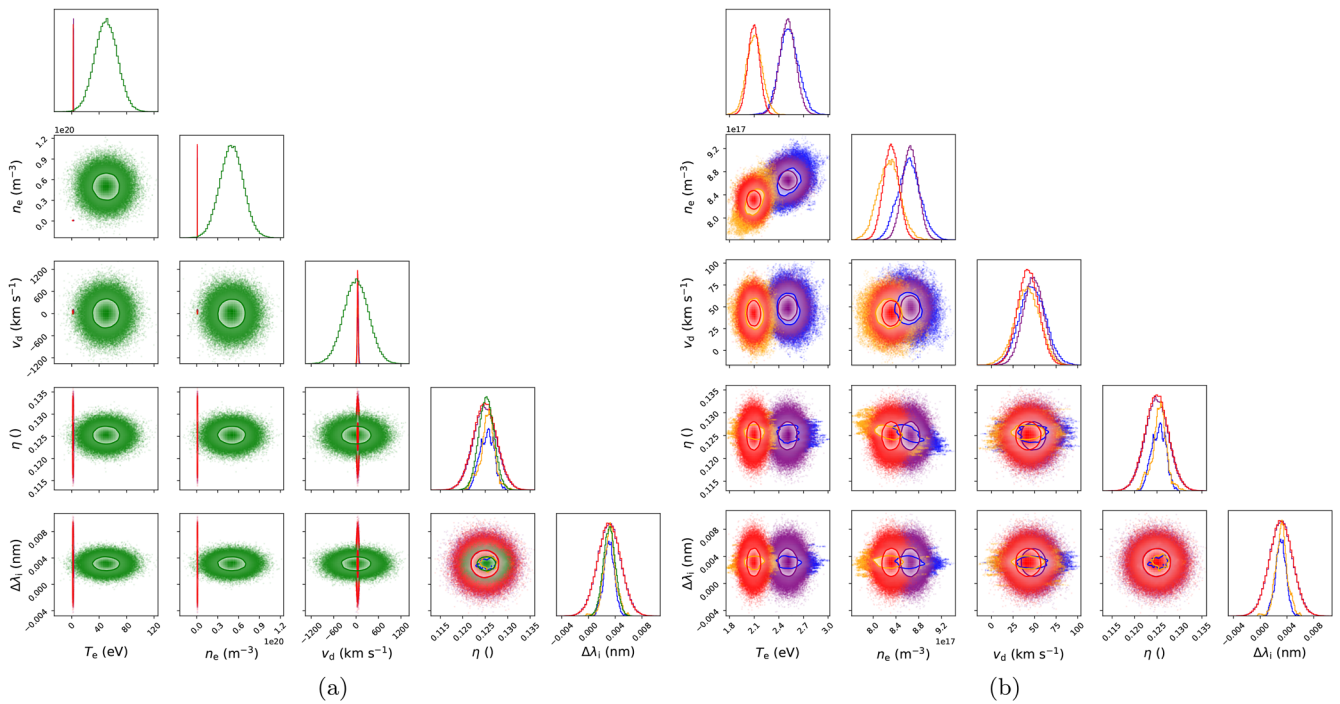


FIG. 15. LTS marginal and joint posterior distribution functions for all LTS model inputs along with the prior and least-squares posterior estimates. (a) Samples from the prior are in green, the Bayesian posteriors for the Maxwellian and Druyvesteyn submodels are in blue and orange, respectively, as well as the posterior estimates of the LS inference are in purple and red, respectively. (b) The same without the prior. For the sake of clarity, each joint posterior displays a single Mahalanobis contour containing the most probable 11.8% of the distribution.

inferences like the one presented here, in which there is an LRS inference informing the nuisance parameters of the LTS inference, is very sensitive to the nuisance parameters of each previous inference.

As the width of the nuisance parameter priors increases, the more divergence between the MAP estimate and the most probable value from the Bayesian inference was observed. When the width of the prior for η in the LS inference was artificially increased by more than 15%, the width of the posterior of the electron number density was observed to extend past 10% compared to the base case. This indicates that for low SNR LRS measurements, or cases like the one presented here where relatively high uncertainties between the estimates of η are present, the LS estimated error bars are not representative of the true uncertainty and that propagation of the uncertainty on η or Bayesian inference is required.

The consistency in predicting a lowermost probable electron number density and electron temperature in the Druyvesteyn cases with similar uncertainty profiles and with all other most probable QoI having almost identical predictions indicates the importance of selecting the best model. Selecting the best model based on the available information and fidelity of the inferences is important as the predicted electron temperatures and densities can vary up to 10% and 15%, respectively.

The $\ln(B_{2,1})$ were calculated using 250 000 samples and were ~ 0 in both cases. This is due to the fact that, as stated in Sec. V A, inaccurate priors lead to indistinguishability of the models. The

values of $\ln(B_{2,1})$ using 250 000 samples and drawing from a uniform prior that was constructed using the minimum and maximum values across both Maxwellian and Druyvesteyn posteriors were calculated to be -1.07 and -35.76 for the 5 and 10 kV cases, respectively. This is consistent with the SNR ratio difference evident in the spectra presented in Fig. 12. This indicates good support and is distinguishable between models in the high SNR case (10 kV), in which the supported model can be taken as ground truth. In the low SNR case, it may be necessary to deploy several other candidate submodels, or bring in additional prior information from other experiments in order to justify the unequal *a priori* probability of the submodels, $\mathbb{P}(M_i) \neq \mathbb{P}(M_j)$, providing additional support of one candidate submodel over the other.

VI. CONCLUSIONS AND OUTLOOK

This paper demonstrates a novel framework for Bayesian model selection in Raman-calibrated laser Thomson scattering. We explore the effects of the accuracy and uncertainty of prior information, the signal SNR, and the model fidelity on model selection in LTS experiments. To do this, we implement a generalized LTS model that, in the limits of $\gamma_{MD} \in [0, 1]$ and a single electron temperature, collapses into a Maxwellian–Druyvesteyn EVDF-based spectrum. The framework can be used to determine that, below SNRs of 5, $\ln(B_{i,j})$ cannot discriminate between candidate submodels, regardless of γ_{MD} , the model fidelity, or the accuracy of prior information. In terms of

model fidelity, as the fidelity of the model increases, or as more than two models are employed, it is clear that *a priori* information on the preference of one model over another is not necessary given the ability of the model selection framework to correctly identify the data generating model through the value of $\ln(B_{i,j})$. However, when expected, this *a priori* information should be factored in to avoid an artificial preference for a Maxwellian EVDF submodel.

In addition, this paper implements a Bayesian inference framework and compares it to posterior estimates from the Laplace approximation, informed via the MAP and uncertainty profiles calculated from the LS inference. The framework demonstrates that at LTS signal SNRs above 10, the Laplace approximation is valid for LRS-calibrated LTS experiments with a high LRS SNR. In cases where the LRS SNR is low, the LTS inference would be further compromised by nuisance parameters from the LRS inference, and a full Bayesian inference or some other form of uncertainty propagation may be required. This validates the widespread use of LS inferences for the determination of the MAP and uncertainty bounds on the QoI, and in cases where the SNR is high, the lengthy implementation of such Bayesian methods is unnecessary. However, in the cases where variable correlations are necessary, the addition of a single nuisance variable has a small but noticeable effect on the predicted correlations for the LTS QoI. Therefore, in cases where the variable correlations are critical, for example, for closure validation in EP-related LTS experiments, full Bayesian inferences may be required depending on the signal SNR.

Within the context of LTS experiments in HET plasmas, specifically for HET closure model validation, it may be important to implement such a Bayesian framework for model selection and parameter inference. The framework presented here can be improved for such an application in several ways. First, a Bayesian inference is implemented for models M_3 and M_4 . This would allow submodels M_i to be compared to M_j for all models i and j to determine the most optimal. The implementation of a plasma that can be independently studied with electrostatic probes and LTS will be of significant benefit for further development of the Bayesian inference and model selection methods. The shortcomings due to leveraging a low order sum of one Maxwellian and one Druyvesteyn brings into question to what extent the sums of Maxwellian and Druyvesteyn distributions can be increased. A future goal of the implementation of this Bayesian framework is to understand the trade-off between extending to higher order sums of Maxwellian and Druyvesteyn distributions and the ability to detect such higher order models as well as the ability to accurately fit all the model parameters.

ACKNOWLEDGMENTS

Jean Luis Suazo Betancourt would like to acknowledge Dr. Adam Friss for assisting with Thomson scattering theory and Dr. Dongho Lee for assisting with data collection.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jean Luis Suazo Betancourt: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (equal); Project administration (lead); Software (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead). **Samuel J. Grauer:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (lead); Writing – original draft (equal); Writing – review & editing (equal). **Junhwi Bak:** Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Adam M. Steinberg:** Funding acquisition (lead); Project administration (lead); Supervision (lead); Writing – review & editing (lead). **Mitchell L. R. Walker:** Funding acquisition (lead); Project administration (lead); Supervision (lead); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from co-authors Mitchell L. R. Walker and Adam M. Steinberg upon reasonable request.

APPENDIX: SCATTERING OF ELECTROMAGNETIC RADIATION

Incident radiation can interact with matter, accelerating charges within the matter, which leads to re-radiation. The efficiency of scattering processes is characterized by a differential scattering cross section that is contingent on the scatterer’s molecule structure. Scattered power intensity is directional, and the distribution of scattered radiation depends upon the reference frame and direction of observation, all of which are wrapped into the differential scattering cross section. Figure 16 depicts the scattering configuration with all relevant angles and vector orientations that appear in the description of our scattering model equations in Appendixes A 1–A 2.

We are interested in the time-averaged scattered power per pulse/shot of incident radiation in photon counts, since our detectors measure (and intensify) each photon that interacts with the photocathode. For detailed derivations, see Refs. 39–44. The general equation representing the spectrally distributed total scattering signal in units of counts per nm is

$$P_\lambda = \underbrace{\eta}_{\text{I}} \underbrace{\frac{\lambda_i}{hc} \Delta\Omega L E_i n}_{\text{II}} \underbrace{\frac{\partial\sigma}{\partial\Omega}}_{\text{III}} S(\lambda), \quad (\text{A1})$$

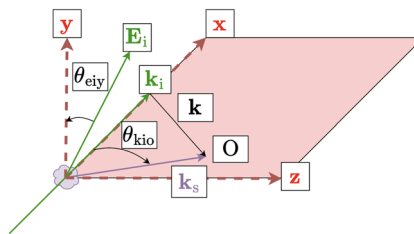


FIG. 16. General wave vector orientations in a scattering experiment.

where η is the intensified photoelectron to scattered photon calibration constant (or the “system calibration constant,” C_{sys}), λ_i/hc gives the photons per joule at the incident wavelength, $\Delta\Omega$ is the solid angle of collection, L is the length of the probe volume, and E_i is the incident laser energy. The braces denote (I) the leading constants that are often grouped as C_{sys} , (II) the total scattered power, and (III) the total scattered power, which is redistributed spectrally due to broadening mechanisms. Also in Eq. (A1), n is the scatterer number density, $\partial\sigma/\partial\Omega$ is the scattering cross section per unit solid angle, and S is the spectral redistribution function. Due to diffraction and the finite nature of the detection system, in particular the spectrometer, the total scattered intensity at any one wavelength is not perfectly captured at that theoretical wavelength. Rather, the aforementioned effects act to redistribute the total scattered intensity around the target wavelength. This is mathematically represented as the spectral distribution function, $S(\cdot)$. The spectral distribution function acts to distribute the total scattered intensity over the domain of interest. The spectral distribution acts as a PDF on the domain of interest and must satisfy the following relationship:

$$\int_0^\infty S(\lambda) d\lambda = \int_0^\infty S(\omega) d\omega = \int_{-\infty}^\infty S(v) dv = 1. \quad (\text{A2})$$

In this expression, ω represents the frequency and v is the velocity. It is paramount that when the scattering equations of interest are cast, the relationship above is satisfied in the domain of interest. If not, the order value and magnitude of the estimated electron number density may be incorrect. Simple dimensional analyses of some of the equations presented in different sources indicate that a wavelength step, $\delta\lambda$, is often included to reconcile the order of magnitude discrepancies that come about from non-normalized spectral distribution functions. The scattering cross section per unit solid angle contains all the physics of the scattering of interest. We discuss the relevant forms to this work, rotational Raman and Thomson scattering, as follows.

1. Rotational Raman scattering

LRS is the inelastic scattering of incident radiation from polyatomic molecules as the result of a net exchange of energy from the incident radiation and the internal energy modes of the molecule.^{39,40} Here, we are concerned with first-order, linear rotational (incoherent) Raman scattering in a neutral diatomic background gas. Closed-form solutions for such LRS are admitted because of the selection rules associated with this type of molecule.^{45,46} For such a molecule, its rotational selection rules for the possible transitions in the rotational quantum number J follow $J' = J \rightarrow J \pm 2$. A Stokes transition is one in which $J \rightarrow J + 2$, and an anti-Stokes transition is one in which $J \rightarrow J - 2$. A transition from $J \rightarrow J$ is a Rayleigh transition and is not considered here. It can be shown^{21,39} that the rotational Raman scattering cross section is the sum of the Stokes and anti-Stokes cross sections as

$$\frac{\partial\sigma^{\text{R}}}{\partial\Omega} = \sum_j \frac{\partial\sigma^{\text{R}}}{\partial\Omega}_{\text{Stokes}} + \sum_j \frac{\partial\sigma^{\text{R}}}{\partial\Omega}_{\text{anti-Stokes}}. \quad (\text{A3})$$

Unlike Thomson scattering, where the polarization of the scattered light is parallel to the polarization of the incident radiation for weakly magnetized plasma in which Faraday rotation is negligible; each individual Raman scattered photon has a polarization that is

random with respect to the incident radiation. Over an ensemble average, the average polarization is quantified through the angle ζ , which is the angle between the polarization of the ensemble-averaged Raman scatter and the incident radiation, as is the case here.

The Stokes and anti-Stokes cross sections are governed by

$$\frac{\partial\sigma}{\partial\Omega_{J'}} = \frac{n_J}{n_g} \{ (1 - \rho) \cos^2(\zeta) [1 - \cos^2(\theta_{\text{kio}}) \sin^2(\theta_{\text{eiy}})] + \rho \} \times \frac{\partial\sigma^\perp}{\partial\Omega_{J'}}, \quad (\text{A4})$$

with ρ , ζ , θ_{kio} , and θ_{eiy} being the depolarization ratio, the angle between the polarization of the scattered Raman radiation and the incident radiation, the angle between the incident radiation propagation direction and the direction of scattering observation, and the angle between the incident radiation polarization and the z axis, in Fig. 16, respectively. Note that the θ for angles should not be confused with the θ representing nuisance parameters in the main sections.

The ratio of the J th population level to the total number of particles is quantified through the partition function as

$$\frac{n_J}{n_g} = \frac{1}{Q_g} g_J (2J + 1) \exp\left[-\frac{\epsilon_J(J)}{k_B T_g}\right]. \quad (\text{A5})$$

The perpendicular Raman cross section, $(\partial\sigma/\partial\Omega)_{J'}^\perp$, is

$$\frac{\partial\sigma^\perp}{\partial\Omega_{J'}} = \frac{64\pi^4}{45} \frac{\gamma^2 b_{J'}}{\epsilon_0^4 \lambda_{J'}^4}. \quad (\text{A6})$$

The anisotropy of the molecular polarizability tensor, γ^2/ϵ_0^2 , is linearly interpolated from experimental measurements at the frequency of interest, as is typical in Refs. 10, 21, and 47.

The rotational energy mode, rotational partition function, ideal gas law, Placzek–Teller coefficients, and Raman frequency shifts are

$$\epsilon_J(J) = hc [B_g J(J + 1) - D_g J^2 (J + 1)^2], \quad (\text{A7})$$

$$Q_g = \frac{(2I_g + 1)^2 k_B T_g}{2B_g hc}, \quad (\text{A8})$$

$$P_g = n_g k_B T_g, \quad (\text{A9})$$

$$b_{J'}(J) = \frac{3}{8} \frac{(2J + 1 \pm 1)(2J + 1 \pm 3)}{(2J + 1)(2J + 1 \pm 2)}, \quad (\text{A10})$$

and

$$\lambda_{J'}(J) = \lambda_i \pm \lambda_i^2 B_g (4J + 2 \pm 4), \quad (\text{A11})$$

respectively. Table I contains the relevant parameter values.

In the case of rotational Raman scattering, we assume that the only relevant source of spectral redistribution is the distribution of the total scattered power due to the finite detection optics, spectrometer, and detection system—commonly referred to as the instrument function—because Doppler and pressure broadening are negligible.

We assume that this (integral normalized) function can be modeled by a Gaussian PDF as

$$S^R(\lambda_{j'}, \lambda_i, \tau) = \frac{1}{\sqrt{2\pi}\sigma_{if}} \exp\left[-\frac{1}{2}\left(\frac{\lambda_i - \lambda_{j'}}{\sigma_{if}}\right)^2\right]. \quad (\text{A12})$$

As mentioned, this redistribution function must adhere to the constraint in Eq. (A2). Traditionally, instrument functions are described with FWHMs, τ . We can relate this to the Gaussian σ_{if} via

$$\sigma_{if} = \frac{\tau}{2\sqrt{2 \ln(2)}}. \quad (\text{A13})$$

In the Raman case, given that we model this as the only source of broadening, the redistribution operation becomes a simple convolution. However, given that each Raman transition produces a single intensity at the particular Raman wavelength, we are mathematically convolving a Gaussian with a Dirac delta function at each transition. As such, the convolution is simply the peak intensity of the Raman line multiplied by the instrument function Gaussian centered at that Raman transition wavelength. Hence, we can explicitly state the expected Raman signal in photon counts per incident laser pulse as

$$P_\lambda^R(\mathbf{x}^R, \boldsymbol{\theta}^R) = \eta \frac{\lambda_i}{hc} \Delta\Omega L E_i n_g(T_g, p_g) \sum_j \frac{\partial\sigma^R}{\partial\Omega_{j'}} S_\lambda^R(\lambda_{j'}, \tau), \quad (\text{A14})$$

in units of counts per nm.

2. Incoherent Thomson scattering

LTS is the elastic electromagnetic scattering of incident radiation from unbounded charged particles and can be coherent or incoherent. Sections 2.2 and 2.4 in Refs. 10 and 47, respectively, discuss the cases in which the experimental setup and plasma conditions meet the conditions for coherent Thomson scattering. In this work, we consider only incoherent LTS.

For LTS, the wavelengths of the scattered radiation are consistent with the Doppler shifted motion of the individual electrons along the scattering wave vector \mathbf{k} ; see Chaps. 1 and 4 of Ref. 43 and Chap. 2 of Ref. 10 for details.

It can be shown that the electron Thomson scattering cross section is

$$\frac{\partial\sigma^T}{\partial\Omega} = r_e^2 [1 - \cos^2(\theta_{kio}) \sin^2(\theta_{eiy})], \quad (\text{A15})$$

with the classical electron radius being $\sim 2.8179 \times 10^{-15}$ m. For LTS, the redistribution of the total scattered power is dominated by two wavelength shifts caused by the relative motion of each scattering electron with respect to the observer of the scatter.⁴³ This is directly linked to the relative velocity of the observer and the scattering electron along the scattering wave vector, $\mathbf{k} \equiv \mathbf{k}_i - \mathbf{k}_s$, with \mathbf{k}_i being the incident propagation wave vector and \mathbf{k}_s being the wave vector along the direction from the scattering volume to the observer. To avoid confusion, the spectral distribution function is explicitly labeled with k to indicate that it is along the direction of the scattering wave vector, in whatever unit vector basis is used to define these vectors. For an ensemble of electrons, this spectral distribution shape function

is directly related to the EVDF along the scattering wave vector, f_k . This distribution happens over tens of nm, an order of magnitude or greater than that of the instrument function broadening, justifying why that contribution to the broadening is often neglected, as is done here. Given this, the total scattered power is redistributed over the electron velocity distribution function along the scattering wave vector. The relationship between the spectral distribution function along the scattering wave vector and the velocity distribution function is $S_k(\lambda) = \frac{1}{k} f_{v_k}(\lambda)$. For a plasma whose electron population is in thermal equilibrium, $S_k(\lambda)$ corresponds to a Maxwellian electron velocity distribution function (EVDF) that can be related to the equilibrium electron temperature T_e . In the case of a non-equilibrium plasma, there are several analytical models that can be applied, including bi-Maxwellian and Druyvesteyn distributions. A brief description of these candidate spectral distribution submodels is presented below. Please note that the spectral distribution functions here implicitly assume that the plasma is low-temperature and neglect the relativistic terms that may be present in some fusion applications.⁴³

3. Maxwellian, Druyvesteyn, and generalized single dimensional electron velocity distribution function submodels

As stated above, electrons adopt a Maxwellian EVDF in the thermal equilibrium case. This is a solution to the Boltzmann equation in which the net effect of any collisional process between like and unlike molecules is zero, regardless of the collision mechanism.⁴⁸ This is encountered in plasmas with high degrees of ionization and collisionality.¹⁰ In this case, the classical electron temperature is the full descriptor for the shape of the distribution. The Druyvesteyn EVDF is a non-equilibrium distribution that describes plasmas where the low degree of ionization and non-constant collision frequency among the electrons lead to a distribution that has a lower density in the high energy regime. In the Druyvesteyn case, the effective electron temperature is defined from a mean energy calculation determined from the moments of the EVDF distribution function.

In the more general case, a cumulative energy distribution function (CEDF) that interpolates between the Maxwellian and Druyvesteyn distributions parameterized by d, T_{eV} and a function of the electron energy in eV^{24,26–28} is

$$F(E_{eV}) = c_1(d) T_{eV}^{-3/2} \sqrt{E_{eV}} \exp\left\{\left[-c_2(d) \left(\frac{E_{eV}}{T_{eV}}\right)^d\right]\right\}, \quad (\text{A16})$$

with c_1 and c_2 being the functions of gamma functions of d .

This CEDF has extensively been used to assess the degree of departure from a Maxwellian distribution in electrostatic probe data, quantified through the parameter d . At $d = 1$, Eq. (A16) is the CEDF of a Maxwellian, and at $d = 2$, it is one of a Druyvesteyn. The conversion between this CEDF and the PDF required in order to cast the spectral distribution function from the eV domain to the nm domain, with the label G representing the generalized distribution, is

$$S_k^G(\lambda) = \frac{\partial E_{eV}}{\partial\lambda} S_k(E_{eV}(\lambda)), \quad (\text{A17})$$

with

$$S_k^G(E_{eV}(\lambda)) = \frac{1}{k} \left[\frac{\partial}{\partial \lambda_s} F \left(E_{eV}(\lambda_s) = \frac{1}{2} m_e v(\lambda)^2 \right) \right], \quad (\text{A18})$$

with, as implicitly stated in the S_k equations above,

$$v = \frac{1}{k} (\omega_i(\lambda_i) - \omega(\lambda) - kv_d^2). \quad (\text{A19})$$

The somewhat arbitrary nature of the functional form of Eq. (A16) (see Ref. 28) and ambiguity in the dimensionality of E_{eV} make solving Eqs. (A17) and (A18) ill-posed. Instead, there is good support in the literature building distributions as weighted sums of other distribution functions, with the most recognizable being the bi-Maxwellian distribution function.^{23–25} To this end, instead of resolving Eqs. (A17) and (A18), a generalized function is defined as a superposition of Maxwellian and Druyvesteyn spectral distribution functions, with

$$P_\lambda^T(\mathbf{x}^T, \boldsymbol{\theta}^T) = \eta \frac{\lambda_i}{hc} \Delta \Omega L E_i n_e \frac{\partial \sigma^T}{\partial \Omega} S_{k,\lambda}^T(\mathbf{x}^T, \lambda_i), \quad (\text{A20})$$

and the spectral distribution function as

$$S_{k,\lambda}^T(\mathbf{x}^T, \boldsymbol{\theta}^T) = \frac{1}{n_e} \left[\sum_{i=1}^{N_M} n_{e,i}^M S_{k,\lambda}^M(n_{e,i}^M, T_{e,i}^M) + \sum_{i=1}^{N_D} n_{e,i}^D S_{k,\lambda}^D(n_{e,i}^D, T_{e,i}^D) \right]. \quad (\text{A21})$$

This function is parameterized by the following QoI and nuisance parameter vectors:

$$\mathbf{x}^T = [\mathbf{T}^M; \mathbf{n}^M; \mathbf{T}^D; \mathbf{n}^D, v_d] \quad (\text{A22})$$

and

$$\boldsymbol{\theta}^T = [\eta, \lambda_i]^T. \quad (\text{A23})$$

Here, \mathbf{n} and \mathbf{T} are the electron number density and temperature vectors. The superscripts indicate a Maxwellian or Druyvesteyn distribution. Using this system, an equilibrium plasma corresponds to the model $N_M = 1$, $N_D = 0$, and $\mathbf{x}^T = [T_e, n_e, v_d]^T$. For simplicity, the superposition is limited to the cases where N_M and N_D are both less than or equal to one. For convenience, we define a distribution fraction for binary Maxwellian–Druyvesteyn plasmas ($N_M = N_D = 1$):

$$\gamma^{\text{MD}} = \frac{n_e^M}{n_e^M + n_e^D} = \frac{n_e^M}{n_e}. \quad (\text{A24})$$

For the Maxwellian and Druyvesteyn distribution functions described above, the spectral distribution functions parameterized by frequency (see Ref. 10 for a full set of derivations) are given by

$$S_{k,\omega}^M = \frac{1}{\sqrt{2\pi}} \frac{1}{k\sigma} \exp \left[-\frac{1}{2} \left(\frac{\omega_i - \omega - kv_d}{k\sigma} \right)^2 \right] \quad (\text{A25})$$

and

$$S_{k,\omega}^D = \beta \frac{1}{k\sigma} \operatorname{erfc} \left[\alpha \left(\frac{\omega_i - \omega - kv_d}{k\sigma} \right)^2 \right]. \quad (\text{A26})$$

The incident and scattered frequencies, the magnitude of the scattering wave vector, and σ are represented by

$$\omega_i = \frac{2\pi c}{n} \frac{1}{\lambda_i} \quad \text{and} \quad \omega = \frac{2\pi c}{n} \frac{1}{\lambda}, \quad (\text{A27})$$

$$k_i = \frac{2\pi}{\omega_i} \quad \text{and} \quad k = \frac{2\pi}{\omega}, \quad (\text{A28})$$

$$k \equiv |\mathbf{k}| = \sqrt{k_i^2 + k_s^2 - 2k_i k_s \cos(\theta_{kio})} \approx 2k_i \sin \left(\frac{\theta_{kio}}{2} \right), \quad (\text{A29})$$

and

$$\sigma \equiv \sqrt{\frac{k_B T_e}{m_e}}. \quad (\text{A30})$$

Finally, as presented, the distributions were derived on the frequency ω domain. To be proper PDFs on the wavelength domain (in nm), a change of variables is required for a proper mathematical redistribution of the scattered power. This is paramount for accurate estimates of the LTS electron number density. In this case, the analytical expression is

$$\frac{\partial \omega}{\partial \lambda} = \frac{2\pi c}{n} \frac{1}{\lambda^2} \quad (\text{A31})$$

such that

$$S_{k,\lambda}(\lambda) = \frac{\partial \omega}{\partial \lambda} S_{k,\omega}[\omega(\lambda)]. \quad (\text{A32})$$

REFERENCES

- ¹P. J. Roberts, B. A. Jorns, and V. H. Chaplin, “Experimental characterization of wave-induced azimuthal ion velocities in a hollow cathode plume,” in *AIAA SCITECH 2022 Forum* (American Institute of Aeronautics and Astronautics, Reston, VA, 2022).
- ²P. J. Roberts and B. Jorns, “Characterization of electron Mach number in a hollow cathode with Thomson scattering,” in *AIAA SCITECH 2023 Forum* (American Institute of Aeronautics and Astronautics, Reston, VA, 2023), p. 1.
- ³B. Vincent, S. Tsikata, and S. Mazouffre, “Incoherent Thomson scattering measurements of electron properties in a conventional and magnetically-shielded Hall thruster,” *Plasma Sources Sci. Technol.* **29**, 035015 (2020).
- ⁴J. P. Boeuf, “Tutorial: Physics and modeling of Hall thrusters,” *J. Appl. Phys.* **121**, 011101 (2017).
- ⁵G. Giono, J. T. Gudmundsson, N. Ivchenko, S. Mazouffre, K. Dannenmayer, D. Loubère, L. Popelier, M. Merino, and G. Olentsenko, “Non-Maxwellian electron energy probability functions in the plume of a SPT-100 Hall thruster,” *Plasma Sources Sci. Technol.* **27**, 015006 (2017).
- ⁶The EVDF is the derivative of the EEPF.
- ⁷K. R. Trent and A. D. Gallimore, “EEDF control of a Hall thruster plasma using a downstream reverse orientation cathode,” in 35th International Electric Propulsion Conference, Atlanta, GA, 2017.
- ⁸C. J. Durot, B. A. Jorns, E. T. Dale, and A. D. Gallimore, “Laser-induced fluorescence measurement of the anomalous collision frequency in a 9-kW magnetically-shielded Hall thruster,” in 35th International Electric Propulsion Conference, Atlanta, GA, 2017.
- ⁹K. Muraoka and A. Kono, “Laser Thomson scattering for low-temperature plasmas,” *J. Phys. D: Appl. Phys.* **44**, 043001 (2011).
- ¹⁰B. Vincent, “Incoherent Thomson scattering investigations in Hall thruster, planar magnetron and ECR ion source plasmas,” Ph.D. thesis (ICARE—Institut de Combustion, Aérothermique, Réactivité et Environnement, Orleans, 2019).

- ¹¹R. L. Washeleski, E. J. Meyer, and L. B. King, "Application of maximum likelihood methods to laser Thomson scattering measurements of low density plasmas," *Rev. Sci. Instrum.* **84**, 105101 (2013).
- ¹²A. Dinklage, R. Fischer, and J. Svensson, "Topics and methods for data validation by means of Bayesian probability theory," *Fusion Sci. Technol.* **46**, 355–364 (2004).
- ¹³R. Fischer, C. Wendland, A. Dinklage, S. Gori, V. Dose, and t. W. A. S. team, "Thomson scattering analysis with the Bayesian probability theory," *Plasma Phys. Controlled Fusion* **44**, 1501–1519 (2002).
- ¹⁴R. Fischer, A. Dinklage, and E. Pasch, "Bayesian modelling of fusion diagnostics," *Plasma Phys. Controlled Fusion* **45**, 1095–1111 (2003).
- ¹⁵S. Kwak, J. Svensson, S. Bozhentkov, J. Flanagan, M. Kempnaers, A. Boboc, and Y. C. Ghim, "Bayesian modelling of Thomson scattering and multichannel interferometer diagnostics using Gaussian processes," *Nucl. Fusion* **60**, 046009 (2020).
- ¹⁶M. M. Saravia, A. Giacobbe, and T. Andreussi, "Bayesian analysis of triple Langmuir probe measurements for the characterization of Hall thruster plasmas," *Rev. Sci. Instrum.* **90**, 023502 (2019).
- ¹⁷N. P. Brown, S. J. Grauer, J. A. Deibel, M. L. R. Walker, and A. M. Steinberg, "Bayesian framework for THz-TDS plasma diagnostics," *Opt. Express* **29**, 4887–4901 (2021).
- ¹⁸C. B. Whittaker, A. Gorodetsky, and B. A. Jorns, "Model inference from electro-spray thruster array tests," in *AIAA SCITECH 2022 Forum* (American Institute of Aeronautics and Astronautics, Reston, VA, 2022).
- ¹⁹J. Bak, J. L. Suazo Betancourt, A. Rekhya, A. Abbasszadehrad, R. B. Miles, C. M. Limbach, and M. L. R. Walker, "High resolution spatially extended 1D laser scattering diagnostics using volume Bragg grating notch filters," *Rev. Sci. Instrum.* **94**, 023003 (2023).
- ²⁰A. C. Tibère-Inglesse, S. D. McGuire, P. Mariotto, and C. O. Laux, "Measurements and analysis of rotational temperatures obtained with Raman and optical emission spectroscopy in a nonequilibrium nitrogen plasma," *Plasma Sources Sci. Technol.* **30**, 125019 (2021).
- ²¹C. M. Penney, R. L. St Peters, and M. Lapp, "Absolute rotational Raman cross sections for N₂, O₂, and CO₂," *J. Opt. Soc. Am.* **64**, 712 (1974).
- ²²A. J. Friss, "Cavity enhanced Thomson scattering for plasma diagnostics," Ph.D. thesis (Colorado State University, 2019).
- ²³H. Sugai, I. Ghanashev, M. Hosokawa, K. Mizuno, K. Nakamura, H. Toyoda, and K. Yamauchi, "Electron energy distribution functions and the influence on fluorocarbon plasma chemistry," *Plasma Sources Sci. Technol.* **10**, 378–385 (2001).
- ²⁴S. F. Adams, J. A. Miles, and V. I. Demidov, "Non-Maxwellian electron energy distribution function in a pulsed plasma modeled with dual effective temperatures," *Phys. Plasmas* **24**, 053508 (2017).
- ²⁵A. A. Haji Abolhassani and J.-P. Matte, "Multi-temperature representation of electron velocity distribution functions. I. Fits to numerical results," *Phys. Plasmas* **19**, 102103 (2012).
- ²⁶J. B. Boffard, R. O. Jung, C. C. Lin, L. E. Aneskavich, and A. E. Wendt, "Optical diagnostics for characterization of electron energy distributions: Argon inductively coupled plasmas," *Plasma Sources Sci. Technol.* **20**, 055006 (2011).
- ²⁷J. T. Gudmundsson and A. Hecimovic, "Foundations of DC plasma sources," *Plasma Sources Sci. Technol.* **26**, 123001 (2017).
- ²⁸H. Amemiya, "Sheath Formation criterion and ion flux for non-Maxwellian plasma," *J. Phys. Soc. Jpn.* **66**, 1335–1338 (1997).
- ²⁹A. Gelman, J. B. B. Carlin, H. S. S. Stern, and D. B. B. Rubin, *Bayesian Data Analysis*, 3rd ed. (Texts in Statistical Science, 2014).
- ³⁰In this text, "distribution" is used loosely to refer to both probability density and mass functions (PDFs and PMFs).
- ³¹R. Fischer and A. Dinklage, "Integrated data analysis of fusion diagnostics by means of the Bayesian probability theory," *Rev. Sci. Instrum.* **75**, 4237–4239 (2004).
- ³²S. J. Grauer, T. A. Sipkens, P. J. Hadwin, and K. J. Daun, "Statistical inversion, uncertainty quantification, and the optimal design of optical experiments," in *Optical Diagnostics for Reacting and Non-Reacting Flows: Theory and Practice* (American Institute of Aeronautics and Astronautics, Inc., Reston, VA, 2023), pp. 1137–1202.
- ³³A. Steinberg and A. K. Patnaik, "Light-matter interactions and their measurement: A primer for optical diagnostics," in *Optical Diagnostics for Reacting and Non-Reacting Flows: Theory and Practice* (American Institute of Aeronautics and Astronautics, Inc., Reston, VA, 2023), pp. 15–74.
- ³⁴N. Friel and J. Wyse, "Estimating the evidence—A review," *Stat. Neerl.* **66**, 288–308 (2012).
- ³⁵M. A. Chilenski, M. Greenwald, Y. Marzouk, J. E. Rice, and A. E. White, "On the importance of model selection when inferring impurity transport coefficient profiles," *Plasma Phys. Controlled Fusion* **61**, 125012 (2019).
- ³⁶F. Sciortino, N. Howard, E. Marmar, T. Odstrcil, N. Cao, R. Dux, A. Hubbard, J. Hughes, J. Irby, Y. Marzouk, L. Milanese, M. Reinke, J. Rice, and P. Rodriguez-Fernandez, "Inference of experimental radial impurity transport on Alcator C-Mod: Bayesian parameter estimation and model selection," *Nucl. Fusion* **60**, 126014 (2020).
- ³⁷J. L. S. Betancourt, J. Lopez-Uricoechea, N. Butler-Craig, A. M. Steinberg, and M. L. R. Walker, "An incoherent Thomson scattering system for plasma boundary studies," *Rev. Sci. Instrum.* **95**, 043001 (2024).
- ³⁸G. A. F. Seber and C. J. Wild, *Nonlinear Regression* (John Wiley & Sons, New Jersey, 2003), Vol. 62, pp. 21–30.
- ³⁹D. A. Long, "Classical theory of Rayleigh and Raman scattering," in *The Raman Effect* (John Wiley & Sons, Ltd, Chichester, UK, 2002), pp. 31–48.
- ⁴⁰N. John and S. George, "Raman spectroscopy," in *Spectroscopic Methods for Nanomaterials Characterization* (Elsevier, 2017), pp. 95–127.
- ⁴¹J. D. Jackson, "Classical electrodynamics," in *Classical Electrodynamics*, 3rd ed. (Wiley, 1998), pp. 661–698.
- ⁴²J. B. Marion and M. A. Heald, "Classical electromagnetic radiation," in *Classical Electromagnetic Radiation* (Dover Publications, 2012) Chap. 10, pp. 335–370.
- ⁴³D. H. Froula, S. H. Glenzer, N. C. Luhmann, and J. Sheffield, "Scattered power spectrum," in *Plasma Scattering of Electromagnetic Radiation* (Elsevier, 2011), pp. 31–44.
- ⁴⁴D. J. Griffiths, *Introduction to Electrodynamics* (American Association of Physics Teachers, 2005).
- ⁴⁵D. W. Ball, "The basics of spectroscopy," in *The Basics of Spectroscopy* (SPIE, 2001), pp. 91–104.
- ⁴⁶J. M. Hollas, "Modern spectroscopy," in *Modern Spectroscopy*, 4th ed. (Wiley, 2004), pp. 103–135.
- ⁴⁷M. J. van de Sande, "Laser scattering on low-temperature plasmas: High resolution and stray light rejection," Ph.D. thesis (Eindhoven Technical University, 2004).
- ⁴⁸W. G. Vincenti, C. H. Kruger, and T. Teichmann, "Introduction to physical gas dynamics," *Phys. Today* **19**(10), 95 (1966).