

# **Machine learning for analysis of atomic spectral data**

M. Cianciosa,<sup>a)</sup> E.H. Martin, and D.L. Green

*Oak Ridge National Laboratory*

K.J.H. Law

*University of Manchester*

(Dated: 20 October 2020)

Physics based forward models are the basis on which many experimental diagnostics are interpreted. For some diagnostics, models can be computationally expensive which precludes their use in real time analysis. Reduced models have the potential to capture sufficient physics thereby enabling the desired real time analysis. Using statistical inference and machine learning techniques the application of a reduced models for inversion of atomic spectral data used to diagnose magnetic fields in a plasma will be examined. Two approaches are considered, a) a reduction of the forward model where traditional inversion can be performed on the proxy model, and b) a reduction of the direct inverse where parameters are a function of measured signal. The resulting inversion is sufficiently fast to be utilized in an online context for digital twinning, and ultimately real-time prediction, design, and control of plasma systems, such as tokamaks. These methods will be demonstrated on both simulated and experimentally measured data.

---

<sup>a)</sup>Electronic mail: [cianciosamr@ornl.gov](mailto:cianciosamr@ornl.gov)

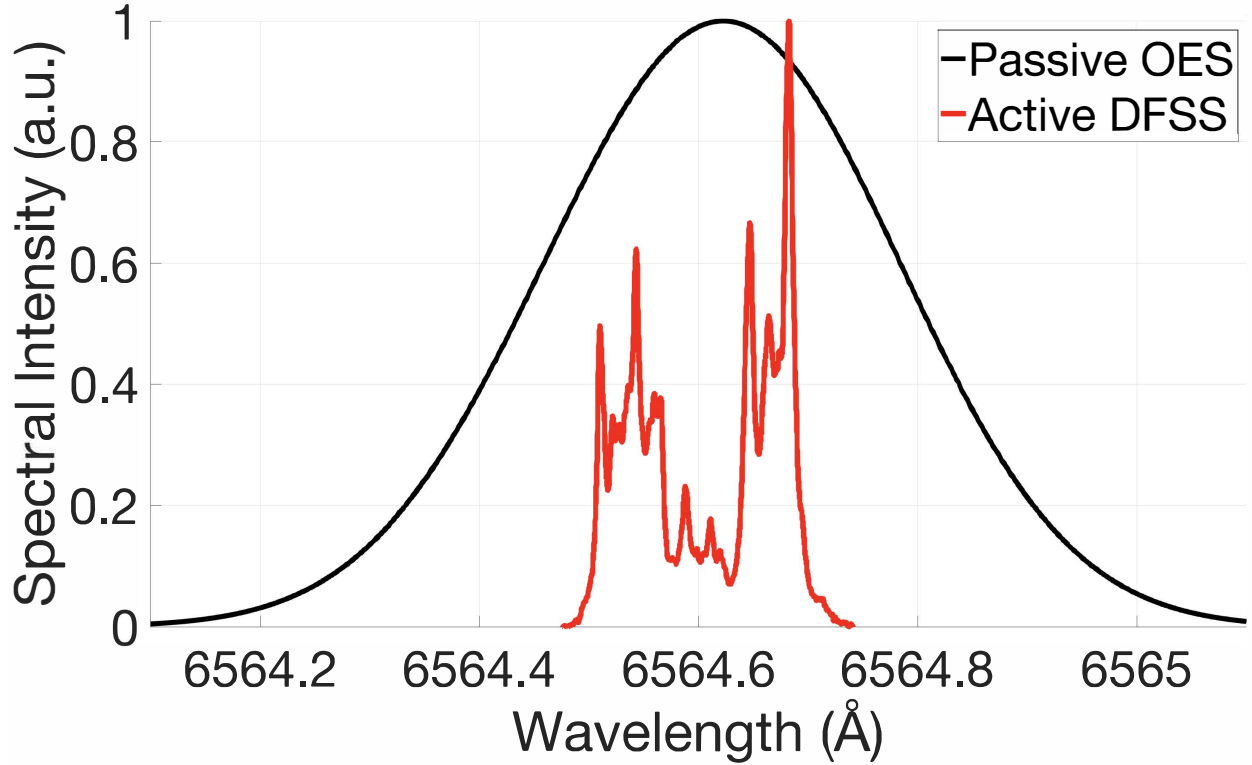


FIG. 1.  $\sigma$  polarized  $H_{\alpha}$  spectra experimentally measured using passive OES and active DFSS experiencing a magnetic field of  $890 \pm 10$  Gauss. The passive measurement results in unresolvable quantum structure for small field amplitudes.

## I. Introduction

Electric and magnetic field vectors are fundamental variables in the equations governing plasma physics, namely the Boltzmann and Maxwell equations. Phenomena such as turbulence, waves, and instabilities can often result in small perturbations of these quantities. These small amplitude fields are of significant importance for the heating, equilibrium, and stability of plasmas. For this reason, a direct, high resolution measurements of these vectors can provide a powerful diagnostic for the investigation of plasma and fusion science.

<sup>1</sup> Notice of Copyright This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

14 Spectroscopy is a common technique used to measure magnetic and electric field vec-  
15 tors through the Zeeman and Stark effect splitting of the spectral line profile. Passive  
16 measurements based on observing light naturally emitted from the plasma have been quite  
17 successful for large amplitude fields ( $|\mathbf{B}| \geq 5000$  Gauss and  $|\mathbf{E}| \geq 500$  V/cm)<sup>1-5</sup>. To mea-  
18 sure small amplitude fields, active methods must be employed to reduce and/or eliminate  
19 spectral broadening mechanisms. Doppler-free saturation spectroscopy (DFSS) is an ac-  
20 tive laser-based technique capable of such measurements<sup>6-9</sup>. This is possible because DFSS  
21 measured spectra can have a resolution approaching the Heisenberg uncertainty principle,  
22 yielding access to the complete quantum structure of the electron. DFSS has been success-  
23 fully demonstrated for measurement of both electric<sup>10-12</sup> and magnetic<sup>13,14</sup> fields in plasmas.

24 In the tokamak fusion reactor, externally driven plasma current in the core is required  
25 to generate a stable magnetic field equilibrium. Radio-frequency (RF) waves are typically  
26 used as a mechanism to drive this current. Recent work has shown that RF wave coupling  
27 with the edge plasma can result in a substantial decrease in the core driven current<sup>15,16</sup>.  
28 Two novel schemes have been designed to optimize RF wave power coupling to the core  
29 plasma<sup>17,18</sup>. Planning is underway to implement DFSS for measurement of the RF wave  
30 electric field vector in the edge plasma. This experimental data will be used to validate  
31 predictive physics models, used to optimize performance, and provide a real-time diagnostic  
32 to monitor operations.

33 Figure 1 presents two  $H_\alpha$  ( $n = 3$  to  $2$ ) spectrum measured under identical conditions in  
34 a magnetized hydrogen/helium plasma using passive optical emission spectroscopy (OES)  
35 and active DFSS. The quantum structure is unresolvable in the passive OES spectra due  
36 to Doppler and instrument broadening. Using active DFSS the Doppler broadening is re-  
37 duced by a factor of 500 and the instrument broadening is effectively eliminated due to  
38 the extremely narrow laser linewidth (100 kHz). The DFSS spectrum conveys the detailed  
39 intricacies of the hydrogen atoms' quantum structure under the effects of a magnetic field.  
40 Quantum mechanical and atomic physics modeling is required to extract the electric and/or  
41 magnetic field vector from DFSS measured spectra. This stipulation is due to the signif-  
42 icant reduction in Doppler broadening and appearance of what is known as the crossover  
43 resonance<sup>19,20</sup>.

44 Ultimately we would like to predict some input parameters  $\mathbf{x}$  of a model  $f(\mathbf{x})$ , from some  
45 diagnostic measurements  $\mathbf{y}$ . Inversion methods utilizing full physics models have drawbacks.

46 Full physics forward models can be computationally expensive. Local optimization methods  
 47 that reduce the number of forward model evaluations used for inversion can take a variable  
 48 number of iterations depending on initial conditions. This non-deterministic behavior makes  
 49 them ill-suited for real-time applications such as feedback control.

50 Machine learning is a collection of methods used to determine unknown functions of  
 51 arbitrary data. In industry, machine learning methods are used to determine functions to  
 52 perform tasks such as identify objects in images, convert spoken words into text, classify  
 53 handwritten digits, or speed up computationally expensive algorithms. Machine learning  
 54 methods such as neural nets have been used to accelerate complex physics models<sup>21</sup>. However  
 55 one of the biggest strengths of machine learning is finding a model when no known model  
 56 exists. In the case of these spectral measurements, there is no known physics based inverse  
 57 model that maps the observable spectrum back to the underlying physical quantities.

58 In this paper we apply machine learning techniques to find simplified forward models and  
 59 an unknown direct inverse of the complex quantum mechanical physics mapping between  
 60 the underlying quantities and the observable DFSS spectrum. Similar techniques of reduced  
 61 forward models have been used to speed up various optimization problems<sup>22-24</sup> while neural  
 62 networks have been used develop direct inverse models for a variety of applications<sup>25,26</sup>. The  
 63 results of the machine learning analysis are compared with spectral fitting conducted using  
 64 a standard global optimization algorithm that searches the entire parameter space<sup>27</sup> of the  
 65 full physics model.

66 This paper is organized as followed. Section II contains a description of the physical  
 67 diagnostic and the full physics model used to interpret measurements. Section III A contains  
 68 a description of the machine learning methods used. Section IV contains a description of  
 69 the training data used. Sections V and VI describes the results of reduced forward and  
 70 inverse models. Section VII benchmarks the performance of the reduced model on actual  
 71 experimental data.

## 72 II. Description of the Diagnostic

73 DFSS is based on exciting electronic transitions using a tunable laser source<sup>6,7</sup>. The  
 74 spectrum is obtained by measuring the absorbed laser power as the laser frequency is swept  
 75 over the transitions of interest. To obtain the Doppler-free resolution, the laser beam is

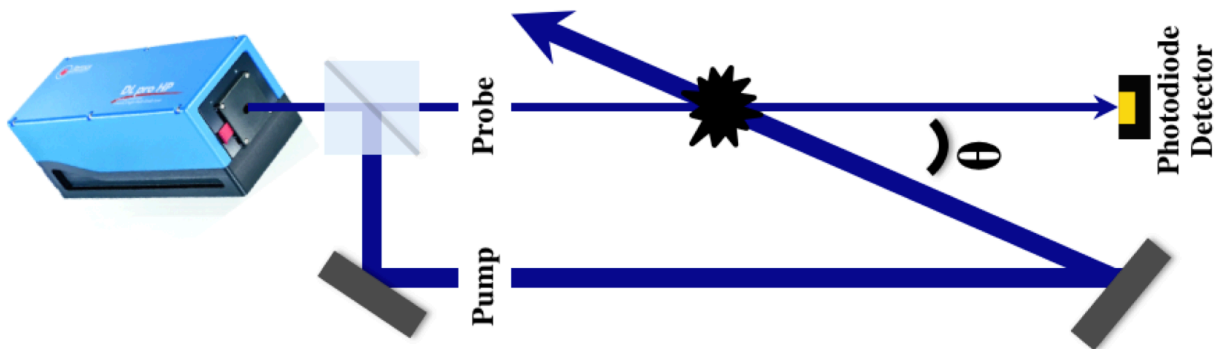


FIG. 2. Basic schematic of DFSS experimental setup. The counter-propagating laser beam geometry results in a substantial reduction in the Doppler broadening present in the measured spectrum.

split (90/10) into two separate beams referred to as the pump and probe. The beams are aligned such that they are counter-propagating at a small angle and overlap at the desired measurement location. The counter propagating geometry allows for the detection of excitation events that occur in atoms having a velocity vector nearly perpendicular to both beams<sup>9</sup>. Accessing a select group of atoms in phase space is possible because the Doppler effect is proportional to the component of the atom's velocity in the direction of the laser beam. The result is an effective reduction in the Doppler broadening associated with the DFSS spectrum. Figure 2 presents a basic schematic of the experimental setup associated with DFSS. The photodiode depicted is used to measure the absorbed laser power and thus the spectrum by sweeping the laser frequency.

DFSS  $H_\alpha$  spectra are modeled using the Explicit Zeeman Stark Spectral Simulator (EZSSS) code<sup>27</sup>. The EZSSS code implements a non-perturbative numerical method to solve the time-dependent Schrodinger equation for the electronic quantum structure of hydrogen and helium-like atom experiencing electric and magnetic field. This method involves decomposing the time dependent Hamiltonian into a Fourier sum. Using the coupled basis wave functions, the individual matrix elements of the Hamiltonian are calculated analytically and the eigenvalue problem is solved resulting in the energies and wave functions of the atom's quantum states. The spectral line profile is calculated using time-dependent first order perturbation theory and the electric dipole connection operator<sup>27</sup>. When considering the simultaneous effects of  $\mathbf{E}$  and  $\mathbf{B}$  fields, the spectral line profile is sensitive to all three

96 components of the vectors.

### 97 **III. Inverse Methods**

98 Measurements of the plasma properties are obtained by fitting the DFSS  $H_\alpha$  model using  
99 the measured spectra through model inversion. Typically a model cannot fit an observation  
100 exactly. Observations can contain measurement noise and errors or models may not fully  
101 describe the observations. The problem of model inversion is an inherently statistical process.

102 In a Bayesian approach<sup>28</sup>, we seek the most probable parameters  $\mathbb{P}(\mathbf{p}|\mathbf{y})$  given some given  
103 some evidence  $\mathbf{y}$ . The optimal parameters  $\mathbf{p}$  are typically unknown a priori  $\mathbb{P}(\mathbf{p}) = 1$ , so the  
104 best fit model is found at the point of maximum likelihood  $\mathbb{P}(\mathbf{p}|\mathbf{y}) \propto \mathbb{P}(\mathbf{y}|\mathbf{p})$ . Assuming a  
105 Gaussian distribution of the likelihood

$$\mathbb{P}(\mathbf{y}|\mathbf{x}) \propto \exp \left[ -\frac{1}{2} \|M(\mathbf{p}) - \mathbf{y}\|^2 \right] \quad (1)$$

106 the problem of finding the optimal parameters reduces to a minimizing a (possibly nonlinear)  
107 least squares objective function

$$\ln \mathbb{P}(\mathbf{p}|\mathbf{y}) \propto \chi^2(\mathbf{p}) = \|M(\mathbf{p}) - \mathbf{y}\|^2. \quad (2)$$

108 The objective functions are the elements of the model we are trying to fit to the observable  
109 quantities. In the case of the DFSS diagnostic, it is the spectral line amplitudes. This  
110 function is minimum when the modeled spectrum  $f(\mathbf{x})$  best matches the measured spectrum  
111  $\mathbf{y}$ .

112 When the number of model parameters is larger than the number of elements in the  
113 objective function, the system is under determined. Under these conditions, multiple solu-  
114 tions exist and the optimization converges to a local minima. When the evidence  $\mathbf{y}$  contains  
115 noise or errors, under determined systems can be fitting to the errors and not modeling the  
116 desired physics. When the number of objective functions is larger than the number of free  
117 parameters, the system is over determined. Under these conditions, multiple solutions are  
118 eliminated however the solution may not be exact. This has an advantage for noisy data as  
119 it smooths out the rapid fluctuations in favor of an average solution.

120 The model parameters at the minimum of the objective have the highest probability  
121 of being true value. Since we assumed a Gaussian distribution of likelihood, the posterior

122  $\mathbb{P}(\mathbf{x}|\mathbf{y})$  also has a Gaussian distribution. To map out this distribution, a best unbiased linear  
 123 estimate<sup>29</sup> can be used to propagate the uncertainty from the evidence to the posterior. By  
 124 defining a Jacobian of our model with respect to the model parameters

$$J_{ij} = \frac{\partial}{\partial p_j} M(\mathbf{p}), \quad (3)$$

125 the uncertainty and correlations of the model parameters are obtained from covariance  
 126 matrix

$$\mathbf{C}_p = [\mathbf{J}^T \mathbf{J}]^{-1}. \quad (4)$$

127 The uncertainty in a model prediction can be made via

$$\boldsymbol{\sigma}(\mathbf{p}) = \mathbf{J}^T(\mathbf{p}) \mathbf{C}_p \mathbf{J}(\mathbf{p}) \quad (5)$$

128 where  $\mathbf{J}(\mathbf{p})$  is the Jacobian evaluated at the predicted point.

129 For linear objective functions the minimum can be solved for directly. For non-linear ob-  
 130 jective functions the minimum must be searched for employing an iterative method typically  
 131 using local gradients. Depending on the method used, optimization methods can get caught  
 132 in local minima or take a variable number of iterations to converge to the minimum.

133 For the DFSS diagnostic, a global optimization method is used to find the best fitting  
 134 spectrum<sup>27</sup>. Using global optimization avoids the problem of a non-deterministic solution  
 135 by finding the global optimum. As a consequence, diagnostic analyses can take several hours  
 136 as a large number of forward model evaluations need to be performed to map out the global  
 137 parameter space.

## 138 A. Surrogate Models

139 To overcome the diagnostic analysis bottle neck, we consider two approaches to accel-  
 140 erating diagnostic inversion. Surrogate models will be used to reduce the forward model  
 141  $\mathbf{y} = f(\mathbf{x})$  that can be rapidly inverted, and find a direct inverse function  $\mathbf{x} = f^{-1}(\mathbf{y})$ . In  
 142 either case a database of training data was generated from the physics based forward model  
 143 by scanning over a range of input parameters. While the offline generation of this training  
 144 data is computationally expensive compared to the direct inversion of the physics based  
 145 model, online inversion of spectra is greatly accelerated.

146 The surrogate models will contain a number of tunable parameters  $\mathbf{p}$ . These unknown  
 147 parameters can be determined using the same maximum likelihood method used to invert the

148 full physics model. However, here we seek the optimal tuning parameters given the training  
 149 data  $\mathbb{P}(\mathbf{p} | \mathbf{x}_t, \mathbf{y}_t)$ . To avoid under determined systems, the number of tunable parameters  
 150 in the surrogate model will be limited to less than the number of elements in the objective  
 151 function.

152 Surrogate models will be developed using two methods. General Linear Models (GLM)  
 153 will be used to find a reduced forward model<sup>30</sup>. GLM modes use a linear combination of  
 154 basis functions

$$\mathbf{y} = f_{\mathbf{p}}(\mathbf{x}) = \boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{p} \quad (6)$$

155 to describe the surrogate function. Since these models are linear, they can be solved directly.  
 156 Neural Net (NN) models<sup>30</sup> will be used to find a direct inverse model. NN models will consist  
 157 of a combination of linear functions

$$\mathbf{y} = \mathbf{w}\mathbf{x} + \mathbf{b}, \quad (7)$$

158 with non-linear activation layers. For the neural networks used here, we use a rectified linear  
 159 unit (ReLU) defined by

$$y_i = ReLU(x_i) = \begin{cases} x_i & : x_i \geq 0 \\ 0 & : x_i < 0 \end{cases}. \quad (8)$$

160 Due to this non-linearity training the NN models to determine the optimal parameters  
 161 of the linear layers  $\mathbf{p} = \mathbf{w} \cup \mathbf{b}$ , requires using a gradient descent based methods. In this  
 162 paper, Neural Network models were implemented and trained using the Mathematica Neural  
 163 Network functions built on top of the MXNet framework. Uncertainty in the trained models  
 164 will be propagated from training data using the method of equation 5.

## 165 B. Sine Example

166 To highlight some of the properties and limitations of surrogate models consider some  
 167 one dimensional data. In figure 3, 100 training data elements (Black points) were collected  
 168 by randomly evaluating a sine function (Gray line) from  $x = \{-3, 3\}$ . This example while  
 169 trivial is instructive to illustrate some of the consequences of applying surrogate models to  
 170 DFSS spectral data. You will note that in one direction  $x \rightarrow y$  the function is single valued  
 171 but multi-valued in the opposite direction  $y \rightarrow x$ . Surrogate models were trained using a  
 172 GLM model and three different NN topologies (Solid blue lines) of increasing complexity.



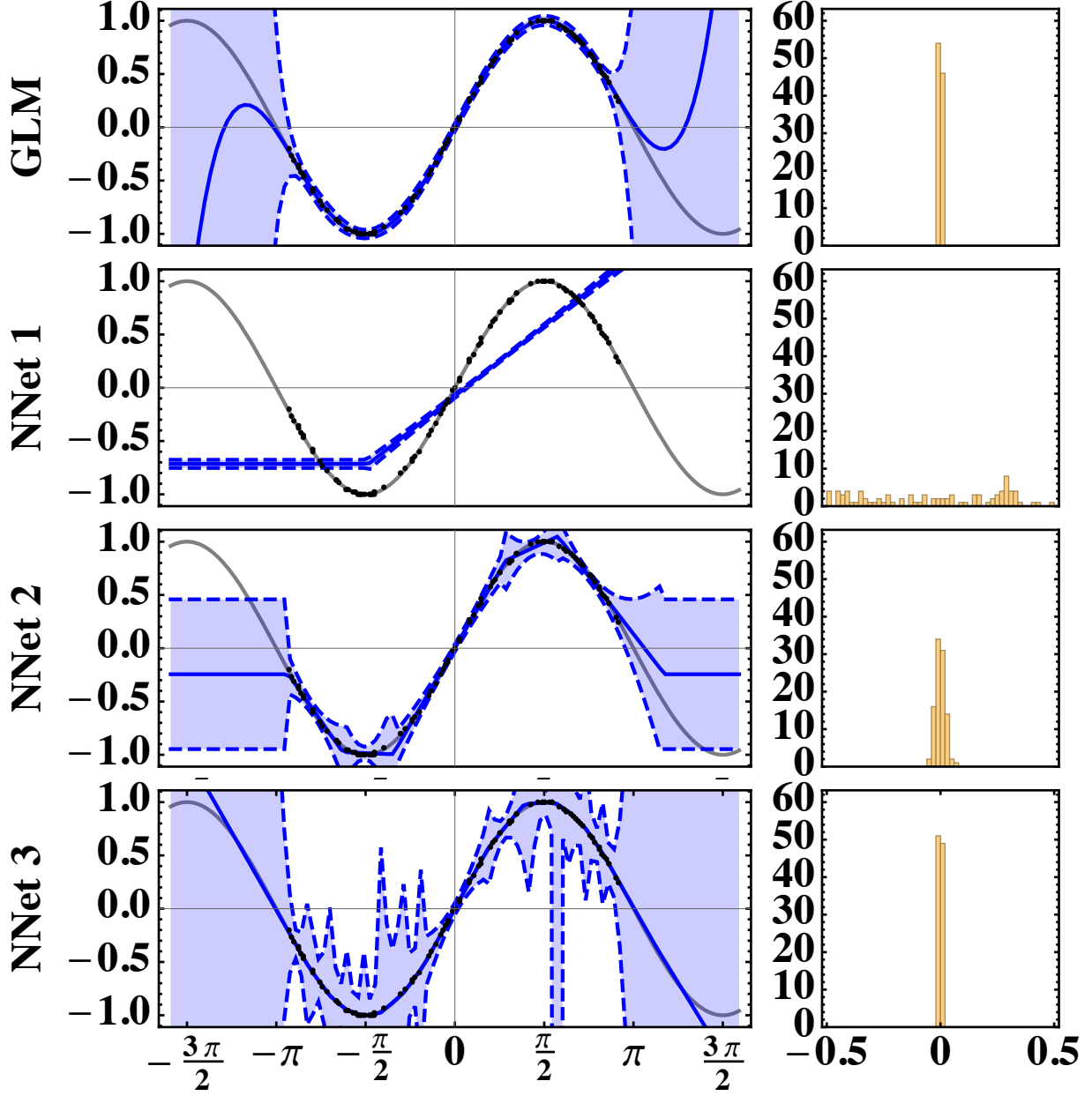


FIG. 3. On the right column, trained ML GLM and Neural Net models (Blue Line) approximating a sine function (Gray Line). Training data (Black Points) was generated from 100 randomly sampled points of the true function from  $x = \{-3, 3\}$ . The GLM model uses a 5<sup>th</sup> order polynomial basis (6 total parameters). NNet 1 uses a single Linear-ReLU layer of width 1 plus output layer (4 total parameters). NNet 2 uses two Linear-ReLU layers of width 3 plus output layer (22 total parameters). NNet 3 uses three Linear-ReLU layers of width 5 plus output layer (76 total parameters). The shaded region represents the model uncertainty propagated from training input data. Beyond the training data, all models quickly diverge from the true function with an accompanied increase in the model uncertainty. In the left column, associated histograms of offset between the true value and machine learned model.

173     Uncertainty in the model prediction is represented by the blue shaded region. The right  
174 column represents the distribution in prediction errors. It should be noted that uncertainty  
175 in the prediction is different from the accuracy of the prediction. For this simple example  
176 case, it is best to view the uncertainty as a proxy representing how much information was  
177 provided to determine the unknown parameters of the function. An ideal surrogate model  
178 would contain both low uncertainty and high accuracy.

179     For the GLM model, 5<sup>th</sup> order polynomial basis functions were used resulting in 6 train-  
180 able parameters. Using a polynomial basis, the trained GLM model finds a solution very  
181 similar to a 5<sup>th</sup> order taylor series expansion of a sine function. Like a taylor series expansion,  
182 the GLM model provides a good approximation within a limited scope but diverges from  
183 the true function at the extrema. As demonstrated dramatically by the model uncertainty,  
184 beyond the bounds of training data, there is no information to determine the model resulting  
185 in a large uncertainty.

186     The simplest neural network (NNet 1) uses two linear layers with a *ReLU* activation  
187 function for total of 4 unknown weight and bias parameters. The other models (NNet 2  
188 and 3) increase in complexity adding an additional, wider linear layers. These two nets use  
189 22 and 76 trainable parameters respectively. Since the neural network models were made  
190 by piecing together linear functions, it should come as no surprise that the trained models  
191 resemble linked linear segments. Since the *ReLU* activation function has a discontinuous  
192 first derivative, the resulting propagated uncertainty is discontinuous as well.

193     (NNet 1, is the equivalent of two lines and woefully inadequate to approximate the true  
194 sine function. The inner layer controls where the *ReLU* function evaluates to a non-zero  
195 value. This activates the slope parameter of the outer layer causing a bend in the model.  
196 As the networks become more complex (NNet 2 and 3), more segments become possible  
197 allowing better linear approximations of the true non-linear function and histograms show a  
198 narrowing in the distribution of model errors. However as the segments become shorter there  
199 are less training data points to determine the parameter values. This results in the increase  
200 in the propagated uncertainty even through the models show a reduction in accuracy.

201     At the ends of the neural net models, there are no more activation functions to break the  
202 lines causing the linear model to diverge from the true non-linear function. For NNet 3, the  
203 ends of the neural net model overlap with a region where the sine function is approximately  
204 linear. This model could still make a fairly accurate prediction beyond the bounds of the

205 training data. However without known data in this region, would be no way to verify the  
 206 result. Thus trained neural net models cannot be relied on beyond their training data.

207 In the second neural network (NNet 2), there the two line segments that do not intersect  
 208 any training data. This would seem to indicate that the optimization method got stuck  
 209 in flat region of parameter space and failed to converge to the global minima. This is a  
 210 possible indication of a non-convex objective function. With different set of initial model  
 211 parameters, the neural net training could have reached a different solution.

212 This example case highlights a number of points to consider when training and evaluating  
 213 surrogate models for the DFSS spectra. While uncertainty will not be propagated due to the  
 214 size of the Jacobian matrix, efforts will be made to reduce the model uncertainty by limiting  
 215 the number of free parameters in the models relative to the amount of training data. Neural  
 216 net models will be trained from multiple initial parameters in case training gets trapped at  
 217 saddle points or other flat regions in parameter space. Model accuracy will be evaluated by  
 218 the lowest overall loss with a narrowest spread in prediction errors.

## 219 IV. Description of Training Data

220 Training data is generated by scanning through the different input parameters ( $|\mathbf{B}|, E_x, E_y, E_z$ )  
 221 of the EZSSS code described in section II. From this code, low resolution data sets are gener-  
 222 ated by sweeping input parameters of magnetic field and electric field vector for five different  
 223 orientations of the laser with respect to the magnetic field direction. Magnetic field is swept  
 224 from 0 Gauss to 1000 Gauss in 100 Gauss steps. Electric field components are swept from  
 225 0 V/m to 200 V/m in 20 V/m steps. Each spectrum is composed of 5000 intensity points  
 226 vs. wavelength with two polarizations. In an effort to reduce the number of unknown  
 227 parameters in the neural net models, the spectra are down sampled. Down sampled spectra  
 228 retain every fifteenth intensity point re-sampling the high fidelity spectra to 334 intensity  
 229 points.

230 Two data sets scan the magnetic field parameter at a higher resolution from 0 Gauss to  
 231 1000 Gauss in 20 Gauss steps while keeping the electric field constant at 0 V/m. Both of these  
 232 higher resolution data sets use the same orientation of the laser with respect to the magnetic  
 233 field direction. One data set differs from the other with an additional physics component  
 234 in the model. One of these data sets was generated using the same simplified model as the

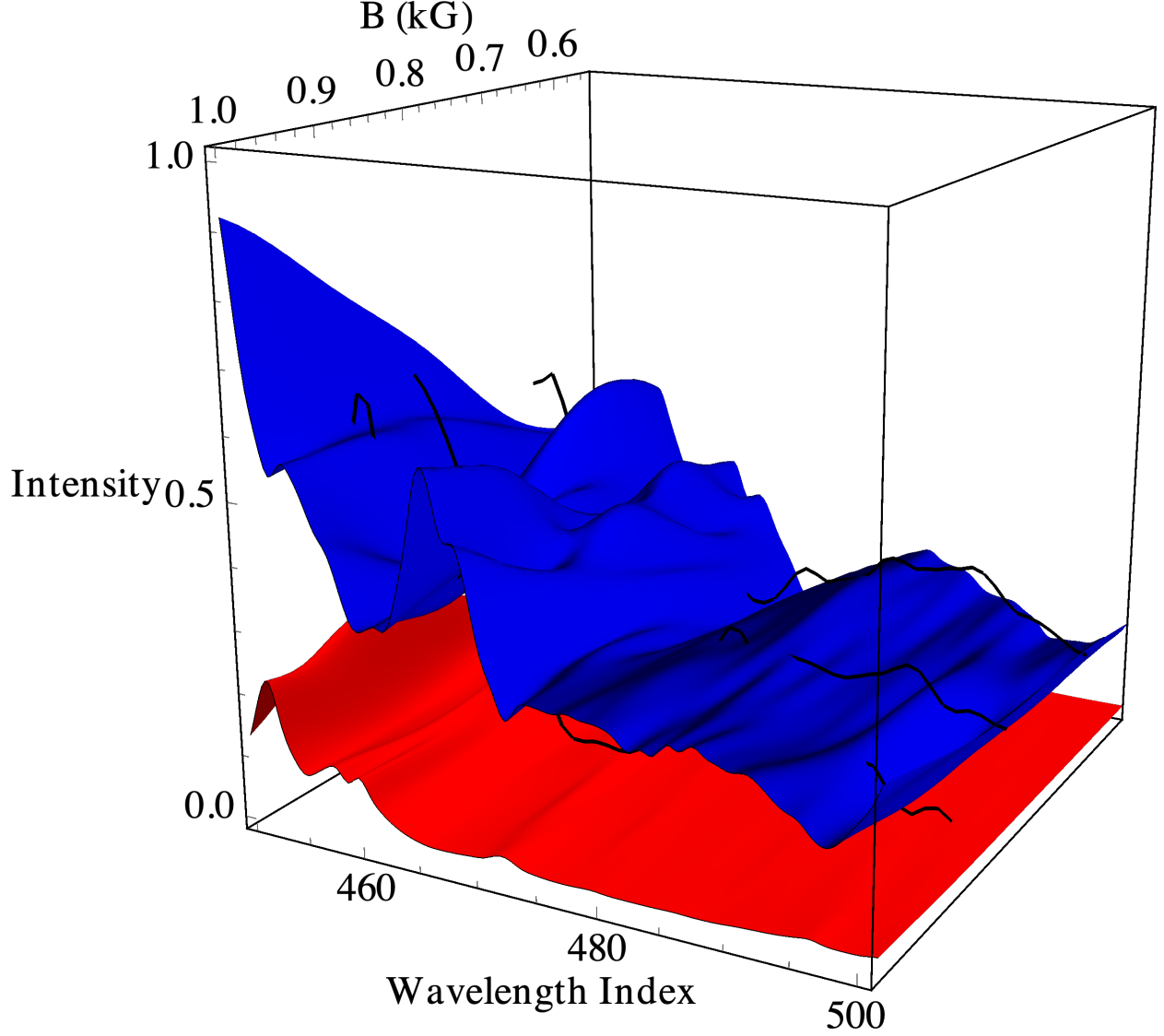


FIG. 4. Zoomed in section of the training data. The blue surface shows the maximum extent of the training data while the red surface shows the minimum extent. The black solid lines are the three experimental spectra. Experimental measured spectra extent beyond the bounds of training data due to addition of spectral lines not present in the training data.

235 aforementioned low resolution four parameter sweeps. The other includes additional physics  
 236 in the form of crossover transitions to model additional sources of emission in the plasma.  
 237 However, the physics based crossover transition model does not accurately capture the true  
 238 plasma response.

239 A final data set consisting of three experimentally measured spectra and an independent  
 240 measurement of the magnetic field strength is provided to test how well the reduced models

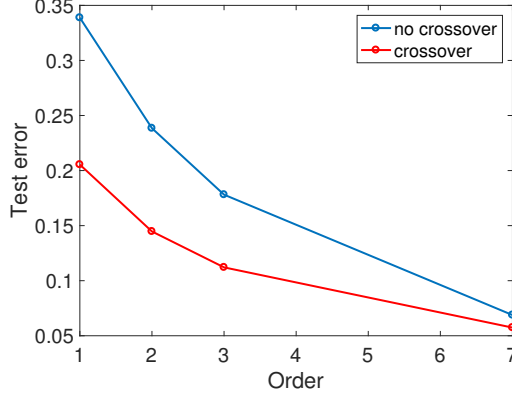


FIG. 5. Reduced model errors over increasing order for the crossover (red) and no crossover (blue) forward model.

perform on real data. The experimental DFSS spectra presented in this paper were measured for  $\mathbf{E}=0$  and  $\mathbf{B}=B_z$ . Measurements were obtained for  $B_z=639, 798$ , and  $890$  Gauss independently measured with a hall effect probe. The error in the known  $B_z$  value is  $\pm 10$  Gauss. The probe laser beam was oriented perpendicular to  $\mathbf{B}$  and the pump/probe angle was set to  $\theta=2^\circ$ . Spectra were obtained having a laser polarization vector perpendicular ( $\sigma$ ) and parallel ( $\pi$ ) to  $\mathbf{B}$ .

Experimentally measured spectra contain contributions due to experimental error, noise and addition physics not captured by the physics model. Figure 4 shows a zoomed in region of the training data. The experimentally measured spectra (Black lines) extent beyond the upper bounds of training data (Blue surface). These experimental data points include points that exists outside of the training data. As shown in section III B it is the hope that the true model scales linearly near the bounds of the training data so that surrogate models can still make close predictions.

For model training, each data set will be randomly split into a training and a validation set. The training set will be used to define the objective function the model should match while optimizing the parameters. The validation set, not seen at all during training, will be used to confirm the accuracy of trained models.

Layer	Input	Width	Num Parameters
Linear	$2 \times 334$	25	16725
ReLU	25	25	0
Linear	25	25	650
ReLU	25	25	0
Linear	25	1 – 4	26 – 104
		Total	17401 – 17479

TABLE I. Direct inverse netural network topology and number of parameters used.

## V. Results of Reduced Forward Model

In this section we examine the results of reduced forward models using the GLM based models. The reduced forward model will be trained using the higher resolution single parameter sweep data for the simplified and crossover transition model. Data sets will be divided in half with one half used for training and the remaining reserved for validation.

Reduced forward models will be obtained from polynomial basis functions  $\psi_k(x) = x^k$  with total degrees of order 1, 2, 3 and 7. These degrees are chosen to examine the effects of increasing order on the reduced models accuracy. It should be noted that for a total polynomial degree of 1, the GLM is exactly the same as a Neural Network with a single linear layer and no activation. Figure 5 shows the relative errors for the simplified and crossover transition models. For both data sets, increasing the total order results in a better recovery of the physics model. Beyond order 7, accuracy does not significantly improve and requires regularization to avoid over fitting.

## VI. Results of Reduced Inverse Model

In this section we train and validate neural network models to make a direct inverse prediction given an input spectrum. Training sets are randomly divided into equal sized training and validation sets. The optimal neural net model will be determined by minimizing the objective function (Equ. 2) with respect to the training data set. Training will continue until the lowest objective function with respect to the validation set is reached. Magnetic

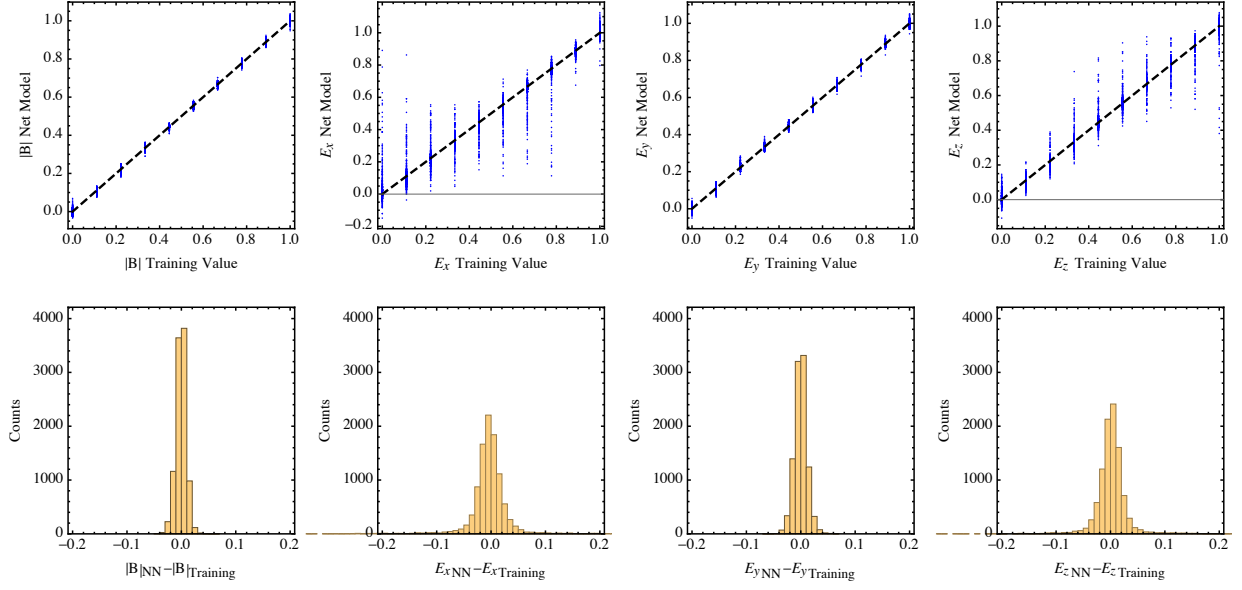


FIG. 6. Results of a neural network model trained to predict the direct inverse of spectral data to all input parameters. The top row compares the model prediction (Blue points) to the idealized perfect prediction (Black dashed line) for each plasma parameter for the full training set. The bottom row histograms show the distribution of prediction errors. The neural net direct inverse model makes a good prediction in magnetic field but poor predictions of electric field vector.

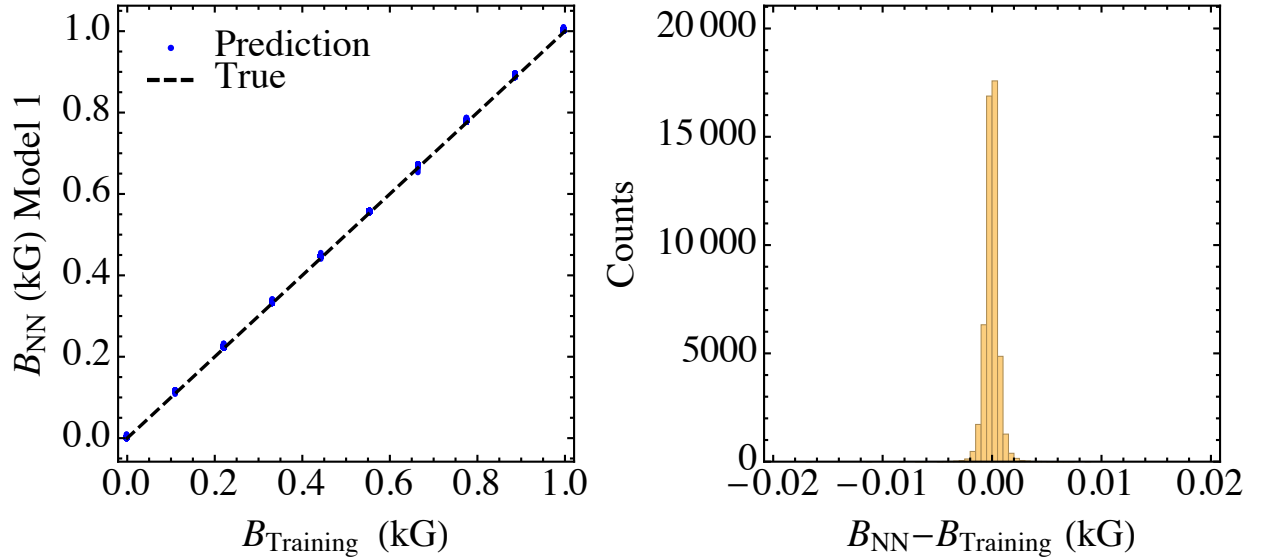


FIG. 7. Results of training the direct inverse neural network model on a combined data set to predict only the magnetic field.

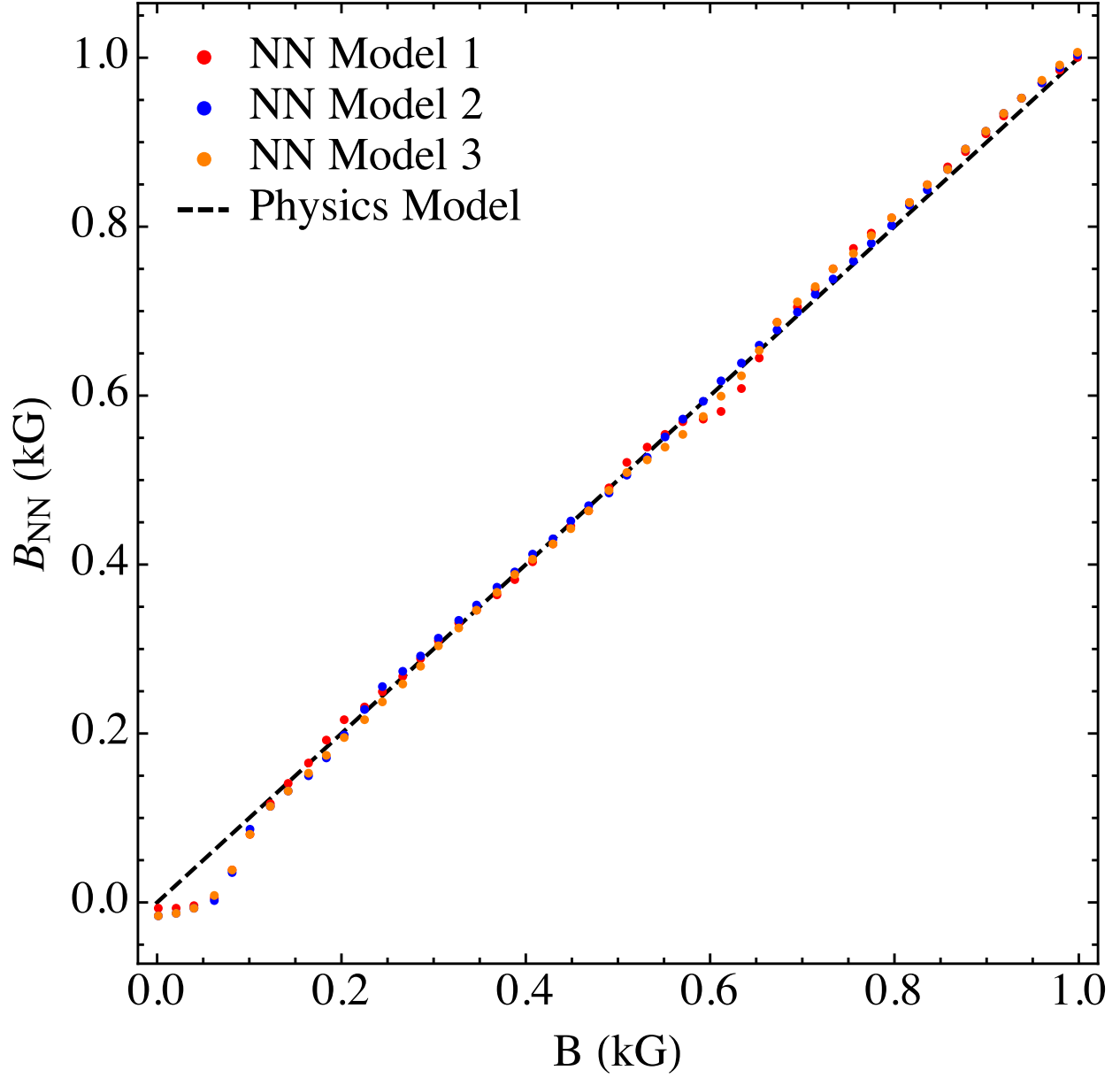


FIG. 8. Validation of low resolution training against high resolution predictions.

and electric field training data is normalized to fall between zero and one.

In an effort to avoid degeneracy, the total number of unknown neural network weight and bias parameters, is restricted to be less than the total number of objective functions in either the training or validation set. The smallest data set used contains ten thousand data pairs. Evenly split, this restricts the maximum number of free parameters to five thousand times the number of outputs. For the networks used here a five layer neural network was used.



283 The network consists of three linear layers with activation functions between them. The  
 284 width of each layer was then chosen such that the total number of network weight and bias  
 285 parameters does not exceed the degeneracy constraint. Table I shows the network topologies  
 286 used and the number of free parameters for that network.

#### 287 *Four parameter model*

288 The low resolution data set contains the spectra associated with the magnetic field  
 289 strength and the vector components of the electric field. Training a neural network to  
 290 predict all four parameters results in four objective functions per data pair. Using a single  
 291 orientation data set, the total number of free parameters allowed in the network are at most  
 292 are twenty thousand. As shown in table I, this is larger than the total number of parameters  
 293 used for the four output network.

294 Figure 6 shows the results of the trained four parameter model for a single orientation.  
 295 The top row shows the scatter in the prediction from the full data set (Blue points) overlaid  
 296 on the ideal perfect prediction (Black dashed line). The bottom row shows a histogram of  
 297 the errors. The large scatter in the  $E_x$  and  $E_z$  components indicate an insensitivity of the  
 298 model to changes in those parameters. However this same model predicts the results  $|B|$   
 299 and  $E_y$  accurately. As shown in section IIIB, this can be the result of not enough layers  
 300 or width in the neural network model. For other orientations, the neural network shows an  
 301 insensitivity to the different electric field components. This indicates a bias in the model  
 302 orientation relative to the electric field vector direction. In all orientations the magnetic  
 303 field magnitude was accurately predicted.

#### 304 *One parameter model*

305 Since the neural network models best predicted the magnetic field magnitude regardless  
 306 orientation, data sets in all five orientations were merged into a single large data set. This  
 307 larger data set is now used to predict the magnetic field only. Since there is only a single  
 308 objective function per spectrum, the number of free parameters is limited to twenty five  
 309 thousand free parameters. As shown in Tbl. I, this is larger than the total number of  
 310 parameters used for the single output network.

For this training set, the forward model is no longer single valued. A single value of the magnetic field maps to multiple spectra since orientation and electric field information is no longer available. However like the sine function example in section III B, the inverse model remains single valued. While multiple spectra map to the same values, a single spectrum can only map to one magnetic field value. Since the training must match multiple spectra to the same magnetic field value, the training will identify features common to the spectra while ignoring the parts that do not matter. To examine the effects of the different splitting between the training and validation sets, the neural network model will be trained three times with different random selections of the training and validation sets (NN Model 1-3).

Figure 7 shows the scatter in the prediction from the full data set (Blue points) overlaid on the ideal perfect prediction (Black dashed line) and histogram of the errors. The model trained here can accurately recover the training and validation sets. Figure 8 depicts the scatter in the prediction from the higher resolution data set (points) overlaid on the ideal perfect prediction (Black dashed line). For each model, the higher resolution data set predicts very close to the ideal except at low field strengths. The disagreement at low field strengths is expected due to the diagnostics decreasing sensitivity in the range of 0 to 50 Gauss. These results show that the neural networks are good at interpolating between training set points.

## VII. Experimental Inference

The ultimate goal of these reduced models is to make a prediction from experimental data. Like the full physics model, the reduced forward model needs to be fitted to observed signal to invert the measured parameter of interest. By contrast, the direct inverse model takes an input spectra and returns the measured parameters without needing to fit the model. Experimental predictions are complicated by the addition of errors, noise, and physics not fully captured in either the physics based or reduced models. These additional features can cause deleterious effects that can throw off a direct inverse model trained on idealized data. The forward based models work around this by fitting the parts of the spectrum it can while retaining large residuals in the fit for the areas it cannot fit.

Figure 9 shows an example of fitting the reduced forward model to an experimentally observed spectra for the simplified and crossover transition models. Neither reduced model

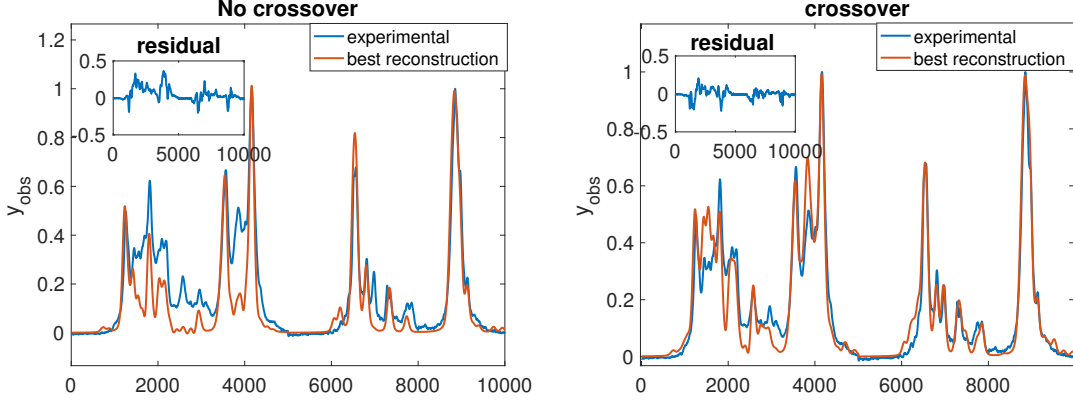


FIG. 9. Comparison of experimentally measured spectra (blue line) with results of the fitted reduced forward model (red line) for reduced models with no crossover transitions left, and crossover transitions right. Both  $\sigma$  and  $\pi$  polarizations are combined into a single spectrum for input into the reduced model. The x axis is the array index of the spectrum intensity.

captures the complete physics in the measurements. However it comes as no surprise that the reduced crossover model captures more of the experimental features with lower total overall residuals since the crossover model is modeling more of the experimental physics.

For the direct inverse models, capturing errors, noise and missing physics requires extrapolating the model beyond the bounds of training data. As shown in section III B, applying machine learning models outside the scope of training data can result in an inaccurate prediction. Since the inverse models were trained on idealized data with known missing physics, we do not expect them to produce accurate predictions. However the predictions while inaccurate, maybe close if the noise, errors and missing physics only represent a small extrapolation from the bounds of training data.

Figure 10 compares the experimental measurements to the reduced forward and direct inverse models and full physics model. The full physics and GLM forward models predict magnetic field strength by inverting the model to match the measured spectra. Inversion of the no crossover physics and GLM forward model show the best prediction of experimental data. Crossover forward models while able to better fit the spectral lines result in a worse prediction. This is a result deficiency of the crossover physics model. The end effects is that the crossover model spectra fit the wrong peaks in the spectra resulting in poor inversion compared to the no crossover models.

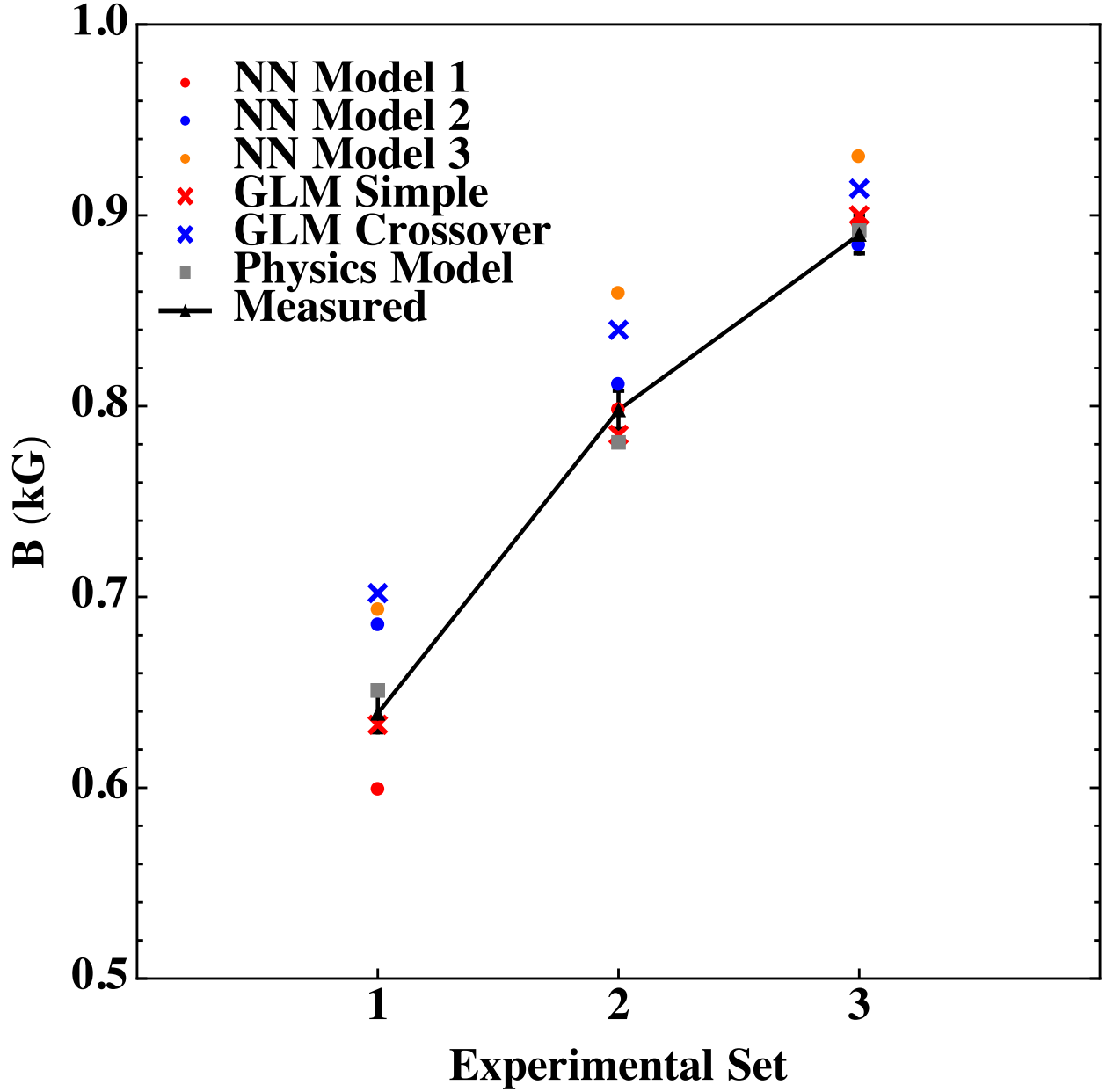


FIG. 10. Reduced model predictions of experimental data.

359 The Neural net direct inverse models predict the physical parameters directly from in-  
 360 put spectra. The addition of spectral lines not accounted for in the idealized training set  
 361 cause the neural network to miss predict the experiment as expected. However they remain  
 362 remarkably close. This provides some promise for the direct inversion method. Provided  
 363 a more accurate physics model, the direct inverse should be able to accurately predict the  
 364 physics parameters of interest.

## 365 **VIII. Summary**

366 Reduced forward models were expected to perform at best as well as inversion of the full  
367 physics model assuming it is coupled to the same inversion method. For any set of conditions  
368 falling within the bounds of the training set, a properly trained reduced model should  
369 accurately interpolate between training data. With the reduced forward model being fast to  
370 evaluate, the speed of inversion is reduced to seconds. However, since the reduced forward  
371 model still needs to be coupled to a inversion method, this method can still potentially  
372 be non-deterministic using local optimization or too computationally intensive using global  
373 optimization techniques.

374 By contrast, the direct inversion method is deterministic in that it will always produce the  
375 same prediction in the same time given an input. Physical parameters can be consistently  
376 predicted in under 2 ms regardless of input spectra. However, the addition of spectral lines  
377 not accounted for within the training set results in the direct inverse needing to extrapolate  
378 beyond the bounds of training data. In order for direct inversion to make an accurate predict  
379 in the presence of noisy data, the model would need the capability to filter out extraneous  
380 line emission. If a spectral wavelength never produces non-zero emission within the training  
381 set, there is no mechanism to set the associated weight parameter of that line to zero during  
382 training. A constraint that minimizes the summed absolute value of the neural network  
383 weights and biases in addition minimizing the loss function could overcome this. However  
384 due to limitations in the functionality of Mathematica's Neural Network functionality, it  
385 was not possible to impose such a constraint. As implemented, the direct inversion model  
386 would need a physics model that could account for the extraneous emission to produce a  
387 more accurate prediction.

## 388 **Acknowledgments**

389 This material is based upon work supported by the U.S. Department of Energy, Office of  
390 Science, Office of Fusion Energy Sciences under award, DE-FG02-04ER54761.

## References

- <sup>1</sup>F. M. Levinton, R. J. Fonck, G. M. Gammel, R. Kaita, H. W. Kugel, E. T. Powell, and D. W. Roberts, *Phys. Rev. Lett.* **63**, 2060 (1989).
- <sup>2</sup>P. G. Carolan, M. J. Forrest, N. J. Peacock, and D. L. Trotman, *Plasma Physics and Controlled Fusion* **27**, 1101 (1985).
- <sup>3</sup>S. Sajjad, H. K. Na, and J. M. Park, *Journal of the Korean Physical Society* **66**, 1687 (2015).
- <sup>4</sup>T. Wujec, H. W. Janus, and W. J. ski, *Journal of Physics D: Applied Physics* **36**, 868 (2003).
- <sup>5</sup>E. Martin, M. Goniche, C. K. adn J. Hillairet, R. Isler, C. Bottereau, L. Colas, A. Ekedahl, S. Panayotis, B. Pegourie, P. Lotte, G. Collendani, J. Caughman, J. Harris, D. Hillis, S. Shannon, F. Clairet, and X. Litaudon, *Plasma Physics and Controlled Fusion* **57**, 065011 (2015).
- <sup>6</sup>V. Letokhov and V. Chebotayev, *Nonlinear Laser Spectroscopy*, 1st ed. (Springer-Verlag, Berlin Heidelberg, 1977).
- <sup>7</sup>K. Shimoda, ed., *High-Resolution Laser Spectroscopy* (Springer-Verlag, Berlin Heidelberg, 1976).
- <sup>8</sup>A. I. Ferguson and J. M. Tolchard, *Contemporary Physics* **28**, 383 (1987), <https://doi.org/10.1080/00107518708224602>.
- <sup>9</sup>E. Martin, A. Zafar, J. Caughman, R. Isler, and G. Bell, *Review of Scientific Instruments* **87**, 11E402 (2016), <https://dx.doi.org/10.1063/1.4961287>.
- <sup>10</sup>M. Adamov, A. Steiger, K. Grutzmacher, and J. Seidel, *Physcial Review A* **75**, 013409 (2007).
- <sup>11</sup>J. Booth, J. Derouard, M. Fadlallah, L. Cabaret, and J. Pinard, *Optics Communications* **132**, 363 (1996).
- <sup>12</sup>S. Nishiyama, K. Katayama, H. Nakano, M. Goto, and K. Sasaki, *Applied Physics Express* **10**, 036101 (2017).
- <sup>13</sup>A. Zafar, E. Martin, and S. Shannon, *Review of Scientific Instruments* **89**, 10D126 (2018), <https://doi.org/10.1063/1.5039334>.
- <sup>14</sup>S. Nishiyama et al., *Journal of Physics: Conference Series* **548**, 012035 (2014).
- <sup>15</sup>I. C. Faust, D. Brunner, B. LaBombard, R. R. Parker, J. L. Terry, D. G. Whyte,

422 S. G. Baek, E. Edlund, A. E. Hubbard, J. W. Hughes, A. Q. Kuang, M. L. Reinke,  
 423 S. Shiraiwa, G. M. Wallace, and J. R. Walk, *Physics of Plasmas* **23**, 056115 (2016),  
 424 <https://doi.org/10.1063/1.4951736>.

425 <sup>16</sup>E. Martin, C. Lau, G. Wallace, S. Shiraiwa, and R. Mumgaard, *Nuclear Fusion* **59**, 076006  
 426 (2019).

427 <sup>17</sup>R. Prater, C. Moeller, R. Pinsker, M. Porkolab, O. Meneghini, and V. Vdovin, *Nuclear*  
 428 *Fusion* **54**, 083024 (2014).

429 <sup>18</sup>P. Bonoli, G. Wallace, S. Shiraiwa, S. Baek, J. Doody, M. Greenwald, B. LaBombard,  
 430 R. Leccacorvi, K. Filar, Y. Lin, E. Marmar, O. Meneghini, T. Palmer, R. Parker, M. Porko-  
 431 lab, B. Sorbom, R. Vieira, A. White, D. Whyte, J. Wright, and S. Wukitch, *Nuclear Fusion*  
 432 **58**, 126032 (2018).

433 <sup>19</sup>L. Magnuir, R. van Bijnen, E. Mese, and R. Scholten, *Journal of Physics B: Molecular*  
 434 *and Optical Physics* **39**, 2709 (2006).

435 <sup>20</sup>E. Saprykin, A. Chernenko, and A. Shalagin, *Optics and Spectroscopy* **113**, 530 (2012).

436 <sup>21</sup>O. Meneghini, C. J. Luna, S. P. Smith, and L. L. Lao, *Physics of Plasmas* **21**, 060702  
 437 (2014), <https://doi.org/10.1063/1.4885343>.

438 <sup>22</sup>A. H. Zaabab, Qi-Jun Zhang, and M. Nakhla, *IEEE Transactions on Microwave Theory*  
 439 *and Techniques* **43**, 1349 (1995).

440 <sup>23</sup>A. Theodoropoulou, R. A. Adomaitis, and E. Zafiriou, *IEEE Transactions on Semicon-*  
 441 *ductor Manufacturing* **11**, 85 (1998).

442 <sup>24</sup>J. S. Han, E. B. Rudnyi, and J. G. Korvink, *Journal of Micromechanics and Microengi-*  
 443 *neering* **15**, 822 (2005).

444 <sup>25</sup>L. Tsang, Z. Chen, S. Oh, R. J. Marks, and A. T. C. Chang, *IEEE Transactions on*  
 445 *Geoscience and Remote Sensing* **30**, 1015 (1992).

446 <sup>26</sup>H. Kabir, Y. Wang, M. Yu, and Q. Zhang, *IEEE Transactions on Microwave Theory and*  
 447 *Techniques* **56**, 867 (2008).

448 <sup>27</sup>E. H. Martin, Electric Field Measurements of the Capacitively Coupled Magnetized RF Sheath Utilizing  
 449 Ph.d. thesis, North Carolina State University (2014).

450 <sup>28</sup>D. S. Sivia and J. Skilling, Data Analysis A Bayesian Tutorial, 2nd ed. (Oxford University  
 451 Press, 2006).

452 <sup>29</sup>C. S. Jones and J. M. Finn, *Nuclear Fusion* **46**, 335 (2006).

453 <sup>30</sup>C. Bishop, Pattern recognition and machine learning, 1st ed. (Springer-Verlag, New York,

454 2006).