Experimentally testing the dependence of momentum transport on second derivatives using Gaussian process regression

Experimentally testing the dependence of momentum transport on second derivatives using Gaussian process regression‡

M A Chilenski[1], M J Greenwald[1], A E Hubbard[1], J W Hughes[1], J P Lee[1], Y M Marzouk[2], J E Rice[1], and A E White[1]

[1] Plasma Science and Fusion Center, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA

[2] Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA

E-mail: markchil@mit.edu

Abstract. It remains an open question to explain the dramatic change in intrinsic rotation induced by slight changes in electron density (White et al. 2013, *Phys. Plasmas* **20**, 056106). One proposed explanation is that momentum transport is sensitive to the *second* derivatives of the temperature and density profiles (Lee et al. 2015, *Plasma Phys. Controlled Fusion* **57**, 125006), but it is widely considered to be impossible to measure these higher derivatives. In this paper, we show that it is possible to estimate second derivatives of electron density and temperature using a nonparametric regression technique known as Gaussian process regression (GPR). This technique avoids over-constraining the fit by not assuming an explicit functional form for the fitted curve. The uncertainties, obtained rigorously using Markov chain Monte Carlo (MCMC) sampling, are small enough that it is reasonable to explore hypotheses which depend on second derivatives. It is found that the differences in the second derivatives of n_e and T_e between the peaked and hollow rotation cases are rather small, suggesting that changes in the second derivatives are not likely to explain the experimental results.

PACS numbers: 02.50.Cw, 02.50.Ey, 02.50.Fz, 02.50.Tt, 02.60.Ed, 02.60.Jh, 02.70.Rr, 02.70.Uu, 07.05.Kf, 52.55.Fa, 52.25.Fi, 52.70.Kz

‡ Preprint submitted to Nuclear Fusion.

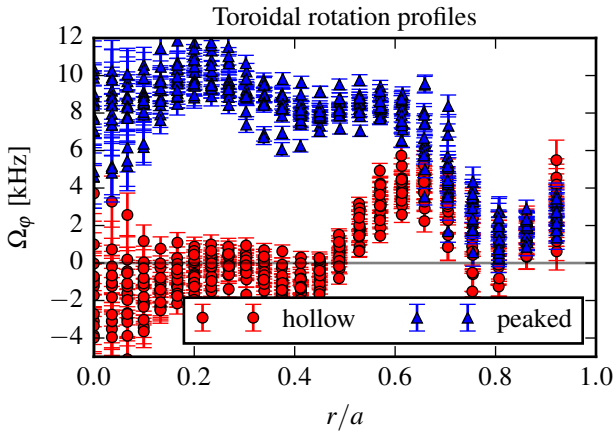

Figure 1: Toroidal rotation profiles for the shots with hollow (red circles) and peaked (blue triangles) rotation profiles. The data shown cover $0.9\,\mathrm{s} \leq t \leq 1.2\,\mathrm{s}$. The hollow rotation shot has n_e approximately 20% higher than the peaked rotation shot, as shown in figure 3.

1. Introduction and motivation: rotation reversals and second derivatives

Explaining the dramatic change of intrinsic rotation with slight changes in density is an open question [1–5]: White et al. [5] present a pair of discharges where the turbulent drive terms are essentially the same within error bars, the only noticeable difference being that the electron density is slightly higher in one discharge. Despite this apparent similarity, the toroidal rotation profile is dramatically different between the two discharges: peaked and strongly co-current in the lower density case and hollow and slightly counter-current on axis in the higher density case. The rotation profiles for these shots are shown in figure 1, the electron density profiles are shown in figure 3 and the electron temperature profiles are shown in figure 4.

Ida et al. [6] and the references therein discuss the experimental connection between the ion temperature gradient and the intrinsic toroidal rotation. Beyond this, there are theoretical reasons to believe that the *second* derivatives of the profiles play a role in momentum transport [7–18]. Barnes gave a very intuitive description of two of the ways this can occur [13]:

(i) The strength of the turbulence depends on the profile gradient. Therefore, the second derivative of the profile corresponds to a gradient in the strength of the turbulence. Consider two ensembles of trapped electrons which start from the same place at the outboard midplane of an up-down symmetric plasma which have opposite parallel velocities. Supposing the toroidal plasma current is in the same direction as the toroidal magnetic field, the electrons for which $v_\parallel > 0$ drift radially outwards when undergoing a banana orbit while the electrons for which $v_\parallel < 0$ drift radially inwards. In the absence of turbulence (and ignoring other collisions), the electrons all eventually return to the same poloidal location, forming a pair of closed

banana orbits. With turbulence of constant strength (i.e., the second derivatives of the profiles are zero) the losses due to turbulence for both populations of electrons are equal, and no net momentum transport occurs. But if there is stronger turbulent transport on one side than on the other then more of the electrons going one direction will be lost and there will be net transport of momentum. Therefore, the second derivatives of the plasma profiles must be non-zero for there to be net momentum transport.

(ii) When there are spatial gradients in the density and temperature profiles there are diamagnetic corrections to the distribution function of the particles which induce a diamagnetic flow. The strength of this flow depends on the first derivatives of the density and temperature profiles. When there is no initial rotation, this diamagnetic flow cancels with the $\boldsymbol{E} \times \boldsymbol{B}$ flow. But there can be different diffusion coefficients for these two flows. So, if the second derivatives of the profiles are non-zero, there will be a gradient in the diamagnetic flow which can cause it to diffuse at a different rate than the $\boldsymbol{E} \times \boldsymbol{B}$ flow, thereby creating a net diffusion of momentum.

Lee et al. [16] puts this in a particularly clear mathematical form: the intrinsic momentum transport is shown to consist of six components (equation (37) of [16]). Of these, the term arising from the different diffusion coefficients for the diamagnetic and $\boldsymbol{E} \times \boldsymbol{B}$ flows, $\Pi_{\text{int}}^{\Delta\chi_\varphi} \propto \partial\Omega_{\varphi,\text{d}}/\partial r$, is found to be a major contributor to the momentum transport in GS2 [19] simulations of the Cyclone Base Case [20, 21]. Because the diamagnetic toroidal rotation frequency $\Omega_{\varphi,\text{d}}$ is, in general, a function of the temperature and density gradients, this is the term which depends on the second derivatives.

In the past it has been widely believed that it is impossible to test hypotheses depending on second derivatives because of how large the uncertainties are expected to grow with each order of derivative, and the risk that an over-constrained fit could mask the true level of uncertainty. In this paper, we show that recent advances in tokamak profile fitting enable this measurement to be made with reasonable uncertainty estimates. It is found that the differences in the second derivative profiles between the peaked and hollow rotation cases are small. This means that, unless the sensitivity to second derivatives is very high, $\Pi_{\text{int}}^{\Delta\chi_\varphi}$ does not appear to be the dominant momentum transport term in the experimental cases, in contrast to what was observed for the Cyclone Base Case. This paper focuses on the experimental measurements and their interpretation; future work will continue the comparison to theory using GS2 simulations of real Alcator C-Mod conditions.

The rest of this paper is organized as follows: section 2 gives a brief review of Gaussian process regression, the profile fitting technique which enables the present work. Section 3 presents the actual fits to the n_e and T_e profiles, and section 4 uses these results to quantify the differences in the second derivative profiles in a statistically rigorous manner. Section 5 summarizes the work and presents the conclusions reached. Appendix A shows how to obtain the derivatives of the squared exponential covariance kernel used for the present work.

2. Estimating derivatives and their uncertainties using Gaussian process regression

Recent advances in tokamak profile fitting using Gaussian process regression (GPR) allow derivatives to be computed along with statistically rigorous uncertainty estimates [22, 23]. Gaussian process regression is a very powerful general-purpose Bayesian nonparametric regression technique [24–31], and the past few years have seen a number of applications of GPR in plasma physics [22, 23, 32–38]. Open-source software is available for both general-purpose GPR [39, 40] and plasma profile analysis [41, 42].

Refer to [27] for an introduction to the general mathematical details, and [22, 23] for an introduction with plasma physics-specific applications. In short, GPR describes the data and the fit using a multivariate normal distribution rather than assuming a specific functional form. It is this lack of assumed functional form which allows GPR to be flexible enough to rigorously quantify the uncertainty in the fitted curve. The posterior distribution from GPR includes the values and derivatives of the profile at all points in space. The properties of the probability distribution are determined by the covariance kernel, $k(x_i, x_j)$, which sets how smooth the profile is: a covariance kernel which decays rapidly with the distance $|x_i - x_j|$ will lead to profiles with fine spatial structure and many wiggles, while a covariance kernel which does not decay rapidly will lead to smooth profiles with little spatial structure. (Here, x_i and x_j are two locations at which predictions and/or observations are to be made.) Because a specific functional form is not assumed, GPR enables a complete accounting of the uncertainty in the profile and its derivatives: the error bars on the fitted curve include all of the flexibility permitted by the data and any prior knowledge regarding the macroscopic smoothness of the profile.

A critical step in fitting profile data using GPR is to infer the hyperparameters[§] $\boldsymbol{\theta}$ which determine how rapidly the covariance kernel decays. This can be handled one of two ways: finding a single estimate for the hyperparameters which are most likely given the data, or drawing samples from the posterior distribution for $\boldsymbol{\theta}$ given the data $\boldsymbol{y}$:

$$\widetilde{\boldsymbol{\theta}} \sim f_{\boldsymbol{\theta}|\boldsymbol{y}}(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{f_{\boldsymbol{y}|\boldsymbol{\theta}}(\boldsymbol{y}|\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta})}{f_{\boldsymbol{y}}(\boldsymbol{y})}, \tag{1}$$

where $\widetilde{\boldsymbol{\theta}}$ is the sample of the hyperparameters, $f_{\boldsymbol{\theta}|\boldsymbol{y}}(\boldsymbol{\theta}|\boldsymbol{y})$ is the posterior density containing everything which is known about the hyperparameters once the data have been measured, $f_{\boldsymbol{y}|\boldsymbol{\theta}}(\boldsymbol{y}|\boldsymbol{\theta})$ is the likelihood function representing the probability of observing the data given a particular value of the hyperparameters, $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is the prior density containing any prior knowledge about typical values for the hyperparameters, and $f_{\boldsymbol{y}}(\boldsymbol{y})$ is a normalization constant. In practice, these samples are drawn using Markov chain Monte Carlo (MCMC) sampling [31, 43, 44], and are then used to estimate

§ The term "hyperparameters" is used to denote the fact that these are the parameters of the prior distribution representing the prior knowledge regarding the smoothness of the profile rather than the parameters of a curve used to explicitly fit the profile. See [27, 31] for more details.

Table 1: Prior distributions used for the hyperparameters of the squared exponential covariance kernel when fitting the rotation reversal data.

Quantity	σ_f	ℓ
n_e	$\mathcal{U}(0, 50\times10^{20}\,\mathrm{m}^{-3})$	$\mathcal{U}(0,15)$
T_e	$\mathcal{U}(0, 50\,\mathrm{keV})$	$\mathcal{U}(0,25)$

the marginalized profile

$$f_{\boldsymbol{y}_*|\boldsymbol{y}}(\boldsymbol{y}_*|\boldsymbol{y}) = \int f_{\boldsymbol{y}_*,\boldsymbol{\theta}|\boldsymbol{y}}(\boldsymbol{y}_*,\boldsymbol{\theta}|\boldsymbol{y})\,\mathrm{d}\boldsymbol{\theta} \tag{2}$$

$$= \int f_{\boldsymbol{y}_*|\boldsymbol{y},\boldsymbol{\theta}}(\boldsymbol{y}_*|\boldsymbol{y},\boldsymbol{\theta}) f_{\boldsymbol{\theta}|\boldsymbol{y}}(\boldsymbol{\theta}|\boldsymbol{y})\,\mathrm{d}\boldsymbol{\theta} \tag{3}$$

and its uncertainty according to the procedure described in [22, 23], where $\boldsymbol{y}_*$ is the predicted profile and/or its derivatives. As was shown in [22, 23], when inferring derivatives it is necessary to use MCMC sampling to obtain a full accounting of uncertainty.

3. Fitting the second derivative profiles of the rotation reversal data

The profile data from the same shots used in [5] were re-fit using GPR, and the second derivatives and their uncertainties were computed. Both of these L-mode discharges had $I_\mathrm{p} = 800\,\mathrm{kA}$, $B_\mathrm{T} = 5.4\,\mathrm{T}$ and were heated with 1.2 MW of ICRF power. For each shot the data were averaged over the period $0.9\,\mathrm{s} \leq t \leq 1.2\,\mathrm{s}$. Data from the core and edge Thomson scattering systems were used for both the n_e and T_e profiles. In addition, the data from two electron cyclotron emission grating polychromator systems were used for the T_e profiles. The equilibrium shape and diagnostic locations for the hollow rotation profile shot are shown in figure 2.

Only the data from inside the LCFS were used, so the rapid change in n_e and T_e at the edge does not need to be fit. This motivates the use of the stationary squared exponential covariance kernel:

$$k_\mathrm{SE}(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{|x_i - x_j|^2}{2\ell^2}\right), \tag{4}$$

where ℓ is the (constant) covariance length scale which sets how fast the correlation drops off and σ_f^2 is the signal variance which sets the extent of variation in the fitted curve. It is very important to note that the covariance length scale ℓ is *not* in any way the same thing as the gradient scale length: even if ℓ is constant throughout the domain, a/L can still vary. The slope at the magnetic axis was forced to zero and very simple uniform prior distributions were used for the hyperparameters, which are given in table 1. In order to ensure a complete accounting of the uncertainty, the hyperparameters were sampled using MCMC using an affine-invariant ensemble sampler [48, 49]. The sampler was run with 200 walkers for 600 samples, 200 samples were burned and the remaining

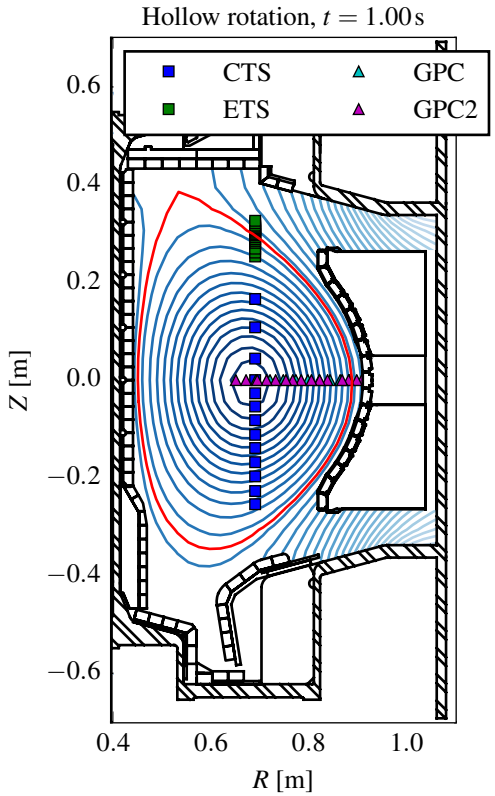

Figure 2: Magnetic equilibrium reconstruction and diagnostic locations for the high density shot with hollow rotation profile. The lower density shot with a peaked rotation profile is essentially the same. In the legend, "CTS" is the core Thomson scattering system, "ETS" is the edge Thomson scattering system, and "GPC" and "GPC2" are the two grating polychromator electron cyclotron emission diagnostics. This figure was produced using eqtools [45–47].

samples were thinned by a factor of 200. The computed profiles are given in figures 3 and 4, summary statistics for the posterior distributions of the hyperparameters are given in table 2 and the posterior distributions themselves are shown in figure 5. The posterior distributions for the hyperparameters are very similar between the two shots for both n_e and T_e.

Figures 3 and 4 show the unnormalized second derivative profiles as well as two possible normalizations:

$$\frac{a}{L_{\nabla n_\mathrm{e}}} = \frac{a\mathrm{d}^2 n_\mathrm{e}/\mathrm{d}r^2}{\mathrm{d}n_\mathrm{e}/\mathrm{d}r} \tag{5}$$

$$\frac{a^2\mathrm{d}^2 n_\mathrm{e}/\mathrm{d}r^2}{n_\mathrm{e}}. \tag{6}$$

The first version was chosen by analogy with the gradient scale length, $a/L_{n_\mathrm{e}} = a(\mathrm{d}n_\mathrm{e}/\mathrm{d}r)/n_\mathrm{e}$, but is made useless in the core by the fact that the first derivative $\mathrm{d}n_\mathrm{e}/\mathrm{d}r$ must go to zero on axis. Therefore, the second version, chosen purely from dimensional analysis, is likely to be more useful in practice.

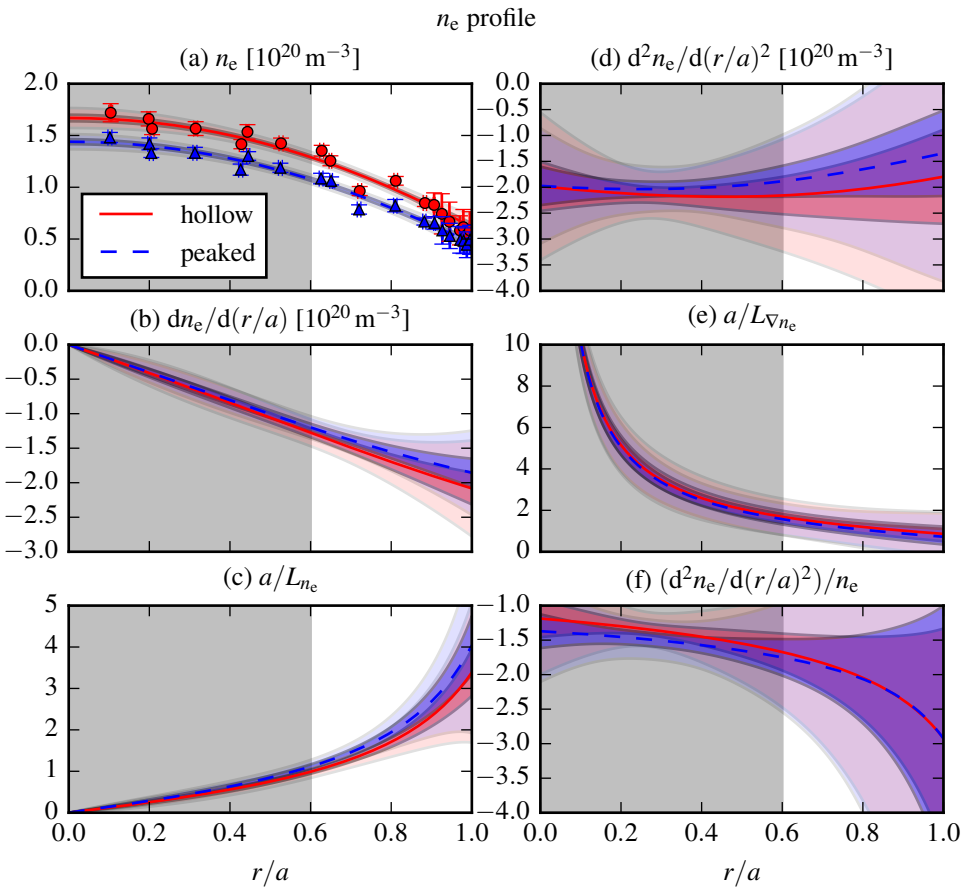

Figure 3: Data and fitted n_e profile for the two shots with different rotation profiles. Red circles and solid curves are from the higher density, hollow rotation profile shot and blue triangles and dashed curves are from the lower density, hollow rotation profile shot. For all curves, the dark band is $\pm 1\sigma$ and the light band is $\pm 3\sigma$. The shaded grey region $r/a < 0.6$ indicates where the largest difference in rotation is. The data were averaged over the same time period as figure 1 before performing the fit. The subplots are: (a) the averaged data and the fitted core profiles, (b) the first derivative of the profile, (c) the normalized inverse gradient scale length of the profile, (d) the second derivative of the profile, (e) the inverse normalized gradient scale length of the gradient (see equation (5)), and (f) the normalized second derivative (see equation (6)).

4. Quantifying the profile differences

In order to test theories which attempt to explain momentum transport in terms of differences in the second derivatives it is necessary to have an objective way of characterizing the difference between two profiles in the presence of very large, overlapping uncertainty envelopes. One approach could be to apply multivariate hypothesis testing as described in [50]. In this case, the test would be deciding whether or not the data are sufficient to reject the null hypothesis that the mean vectors are equal. This oversimplified choice between "mean vectors match" and "mean vectors do

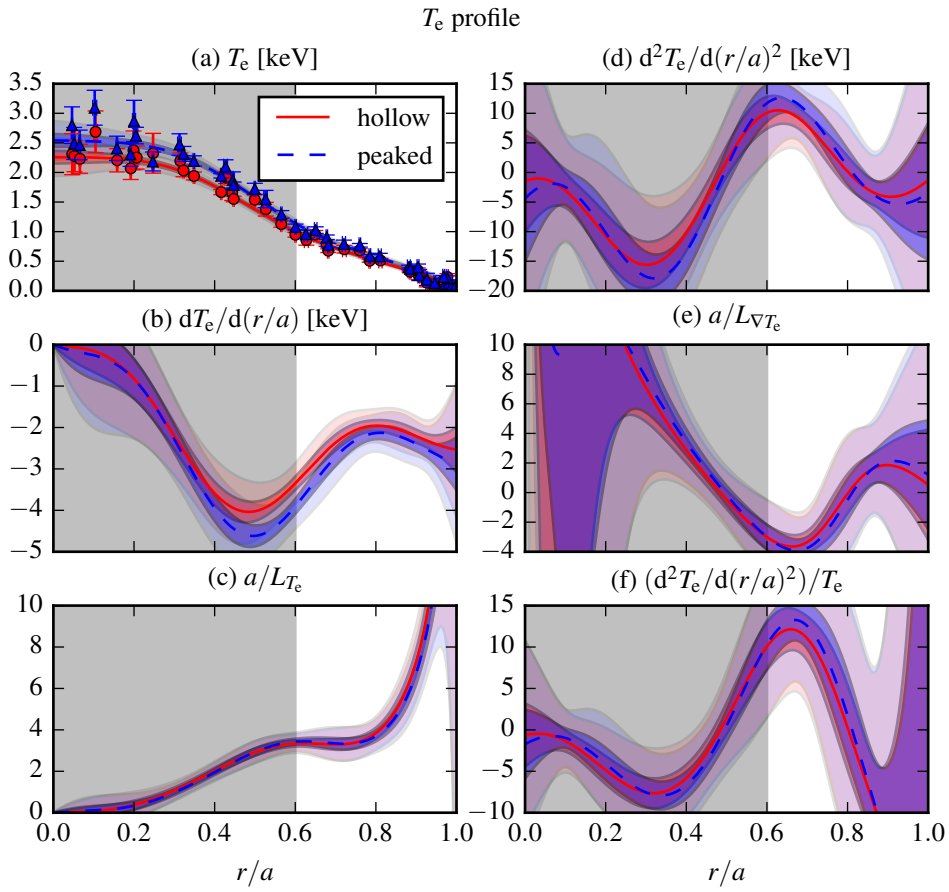

Figure 4: Data and fitted T_e profile for the two shots with different rotation profiles, presented as in figure 3. With the high resolution ECE measurements, the T_e profiles have a shorter covariance length scale and smaller relative uncertainties than the n_e profiles.

Table 2: Summary statistics of the posterior distributions for the hyperparameters.

Quantity	Case	Parameter [units]	Mode	Mean	95% interval
n_e	hollow	σ_f [10^{20} m^{-3}]	1.35	13.6	[1.07 , 45.9]
		ℓ	1.23	3.10	[0.929, 7.19]
	peaked	σ_f [10^{20} m^{-3}]	1.09	13.1	[0.854, 46.1]
		ℓ	1.18	3.13	[0.913, 7.40]
T_e	hollow	σ_f [keV]	1.20	1.92	[0.835, 5.16]
		ℓ	0.317	0.355	[0.255, 0.510]
	peaked	σ_f [keV]	1.33	2.09	[0.933, 5.17]
		ℓ	0.313	0.344	[0.247, 0.477]

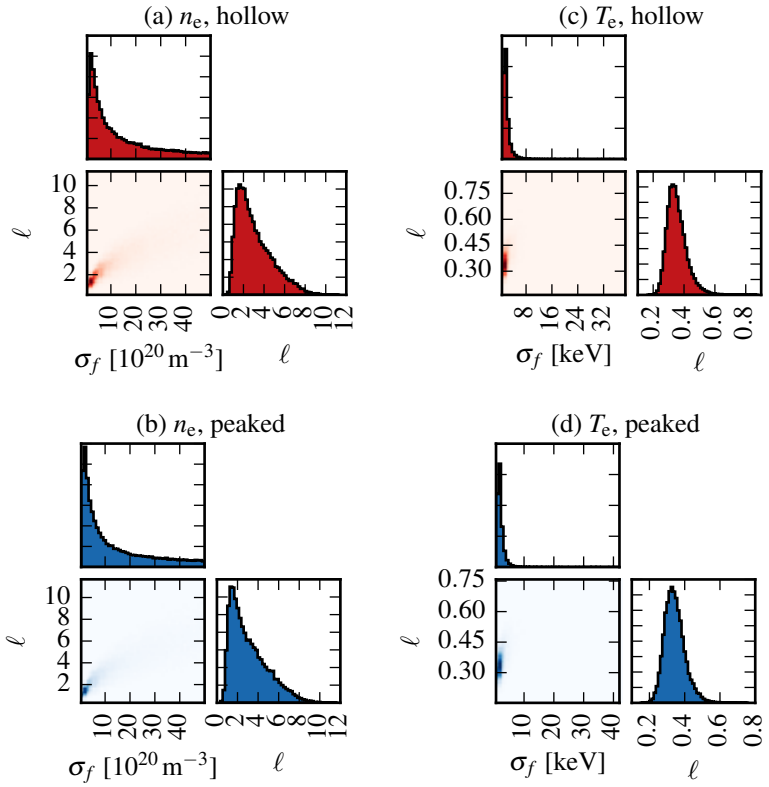

Figure 5: Posterior distributions for the hyperparameters σ_f and ℓ of the various fits. In each subfigure, the lower left corner shows the bivariate histogram of the MCMC samples and the other two plots show the univariate histograms. The posterior distributions for the peaked and hollow cases are essentially the same. Most of the probability mass is concentrated on the modes at $\sigma_f \approx 10^{20}\,\mathrm{m}^{-3}$, $\ell \approx 1$ for n_e and $\sigma_f \approx 1\,\mathrm{keV}$, $\ell \approx 0.3$ for T_e. The posterior distributions are strongly peaked compared to the uniform prior distributions used, which indicates that the sensitivity of the result to the prior distribution is low.

not match" has been noted as a key shortcoming of frequentist hypothesis testing [31]: what really matters is the probability of the profiles being different *enough* to cause a significant change in the rotation profile. The Bayesian framework allows a more useful test to be made: we can actually use the results of the GPR fit to quantify how likely it is that the profiles differ by a given amount.

For a given pair of quantities to compare y_h, y_p (with corresponding realizations of the hyperparameters $\boldsymbol{\theta}_p$, $\boldsymbol{\theta}_h$) we have[‖] $y_h|\boldsymbol{\theta}_h \sim \mathcal{N}(\mu_h, \sigma_h^2)$ and $y_p|\boldsymbol{\theta}_p \sim \mathcal{N}(\mu_p, \sigma_p^2)$. In this context, y_h could be the second derivative at $r/a = 0.5$ of the hollow rotation case and y_p could be the second derivative at $r/a = 0.5$ of the peaked rotation case. The probability

‖ Note that the normalized forms of the second derivative are actually the *ratio* of two Gaussian-distributed random variables. As was shown in [22], as long as the denominator is sufficiently far from zero the result can be approximated as being Gaussian.

density function (PDF) for the difference of these two quantities, $\Delta y = y_\mathrm{h} - y_\mathrm{p}$, is then

$$f_{\Delta y|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}}(\Delta y|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}) = \mathcal{N}(\Delta y; \mu_\mathrm{h} - \mu_\mathrm{p}, \sigma_\mathrm{h}^2 + \sigma_\mathrm{p}^2). \tag{7}$$

Likewise, the cumulative distribution function (CDF) is

$$\begin{aligned} F_{\Delta y|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}}(c|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}) &= \mathbb{P}(\Delta y \leq c|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}) \\ &= \Phi\left(\frac{c - (\mu_\mathrm{h} - \mu_\mathrm{p})}{\sigma_\mathrm{h}^2 + \sigma_\mathrm{p}^2}\right), \end{aligned} \tag{8}$$

where $\Phi(x)$ is the CDF of the standard normal distribution $\mathcal{N}(0,1)$. What is actually of interest are the PDF and CDF averaged over all values of $\boldsymbol{\theta}_\mathrm{h}$ and $\boldsymbol{\theta}_\mathrm{p}$:

$$f_{\Delta y}(\Delta y) = \int f_{\Delta y|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}}(\Delta y|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}) f_{\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}}(\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p})\,\mathrm{d}\boldsymbol{\theta}_\mathrm{h}\,\mathrm{d}\boldsymbol{\theta}_\mathrm{p} \tag{9}$$

$$F_{\Delta y}(c) = \int F_{\Delta y|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}}(c|\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}) f_{\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p}}(\boldsymbol{\theta}_\mathrm{h},\boldsymbol{\theta}_\mathrm{p})\,\mathrm{d}\boldsymbol{\theta}_\mathrm{h}\,\mathrm{d}\boldsymbol{\theta}_\mathrm{p}. \tag{10}$$

These two functions are estimated from the MCMC samples $\{\boldsymbol{\theta}_\mathrm{h}^{(i)}\}$, $\{\boldsymbol{\theta}_\mathrm{p}^{(i)}\}$ according to

$$f_{\Delta y}(\Delta y) \approx \frac{1}{n}\sum_{i=1}^{n} f_{\Delta y|\boldsymbol{\theta}_\mathrm{h}^{(i)},\boldsymbol{\theta}_\mathrm{p}^{(i)}}(\Delta y|\boldsymbol{\theta}_\mathrm{h}^{(i)},\boldsymbol{\theta}_\mathrm{p}^{(i)}) \tag{11}$$

$$F_{\Delta y}(c) \approx \frac{1}{n}\sum_{i=1}^{n} F_{\Delta y|\boldsymbol{\theta}_\mathrm{h}^{(i)},\boldsymbol{\theta}_\mathrm{p}^{(i)}}(c|\boldsymbol{\theta}_\mathrm{h}^{(i)},\boldsymbol{\theta}_\mathrm{p}^{(i)}). \tag{12}$$

These two functions can then be used to compute the posterior intervals which show how large the differences are likely to be.

Figure 6 shows the marginalized PDFs for the normalized and unnormalized second derivatives of n_e and T_e at $r/a = 0.5$. In all six cases the 68% equal-tailed posterior density interval contains zero, suggesting that there is very little difference in both the normalized and unnormalized second derivatives of both n_e and T_e between these two cases. This implies that, unless there is a very strong sensitivity to second derivatives, the momentum diffusion term $\Pi_\mathrm{int}^{\Delta\chi_\varphi}$ cannot account for the change in rotation between these two discharges. This is in contrast to GS2 predictions for the Cyclone Base Case, which suggest that $\Pi_\mathrm{int}^{\Delta\chi_\varphi}$ is one of the dominant momentum flux components [16]. That being said, table 3 shows that there are key differences between the Cyclone Base Case and the real experimental conditions. In particular, the Cyclone Base Case is constructed to be far from marginal stability so as to exhibit robust ion temperature gradient (ITG) turbulence [51]. The results presented here suggest that GS2 simulations of the real experimental conditions should be performed to check whether or not the momentum diffusion term is still expected to be dominant in this pair of discharges.

5. Summary and conclusions

This paper has presented the use of Gaussian process regression (GPR) to infer *second* derivative profiles from noisy, discrete profile data. A complete accounting of uncertainty was obtained through use of Markov chain Monte Carlo (MCMC) sampling. This

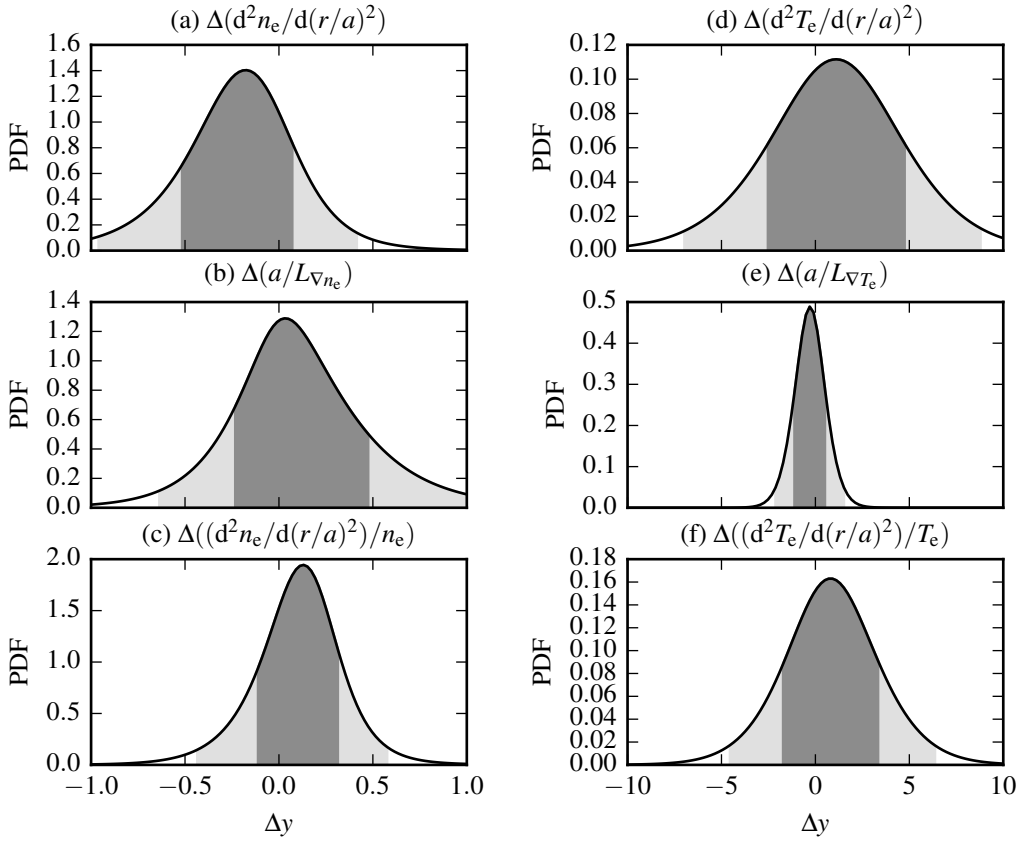

Figure 6: Probability density functions for the differences in the normalized and unnormalized second derivatives of n_e and T_e at $r/a = 0.5$ between the hollow and peaked rotation profile shots. The dark and light shaded regions are the 68% and 95% equal-tailed posterior intervals, which indicate the regions where the differences are most likely to lie. The PDFs are asymmetric as a result of having marginalized over the hyperparameters.

work shows that the uncertainties on second derivatives are not as large as might be feared, and hence we can realistically attempt to test hypotheses which depend on these quantities. This technique was used to test the hypothesis that changes in the second derivatives can explain the dramatic change in intrinsic rotation induced by slight changes in density. The present analysis indicates that there is not a substantial difference between the second derivatives of n_e and T_e in the higher density, hollow rotation and lower density, peaked rotation shots. Therefore, unless a theory is put forth which predicts an exceptionally strong sensitivity to the second derivatives, another explanation is needed.

Acknowledgments

The authors are very grateful for the clear description of the relationship between second derivatives and momentum transport provided by M. Barnes. This material is based

Table 3: Comparison of the simulation parameters for the Cyclone Base Case [21] and the peaked rotation discharge (low power shot from [52]).

Parameter	Cyclone Base Case	Peaked rotation
geometry	shifted-circle model	Miller geometry
electrons	adiabatic	drift-kinetic
R/L_{T_i}	6.9	6.0
R/L_{T_e}	N/A	8.1
R/L_{n_e}	2.2	2.3
r/R	0.18	0.16
q	1.4	1.3
$\hat{s}$	0.8	0.91
$T_\mathrm{i}/T_\mathrm{e}$	1.0	0.89
Z_eff	1.0	1.9
β	0.0	0.0
ν	0.0	0.11

upon work conducted using the Alcator C-Mod tokamak, a DOE Office of Science user facility. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences under Award Number DE-FC02-99ER54512. This material is based upon work supported in part by the U.S. Department of Energy Office of Science Graduate Research Fellowship Program (DOE SCGF), made possible in part by the American Recovery and Reinvestment Act of 2009, administered by ORISE-ORAU under contract number DE-AC05-06OR23100.

Appendix A. Derivatives of the squared exponential covariance kernel

This section obtains an expression for the n^{th} derivative of the D-dimensional squared exponential (SE) covariance kernel. The $n = 2$, $D = 1$ case is what was used for the present work. For this section, $\boldsymbol{x}_i \in \mathbb{R}^D$ is the location an observation or prediction y_i is to be made at, while in the rest of this paper $x_i = r_i/a \in \mathbb{R}$ was a single coordinate.

Recall that the covariance kernel is defined as $k(\boldsymbol{x}_i, \boldsymbol{x}_j) \equiv \mathrm{cov}[y(\boldsymbol{x}_i), y(\boldsymbol{x}_j)]$. Next, note that the relationship between a Gaussian process and its derivatives is

$$\mathrm{cov}\left[y_i,\, \frac{\partial y_j}{\partial x_{j_d}}\right] = \frac{\partial k(\boldsymbol{x}_i,\, \boldsymbol{x}_j)}{\partial x_{j_d}} \tag{A.1}$$

$$\mathrm{cov}\left[\frac{\partial y_i}{\partial x_{i_c}},\, \frac{\partial y_j}{\partial x_{j_d}}\right] = \frac{\partial^2 k(\boldsymbol{x}_i,\, \boldsymbol{x}_j)}{\partial x_{i_c}\, \partial x_{j_d}}, \tag{A.2}$$

where the notation $\partial/\partial x_{j_d}$ refers to a derivative with respect to the d^{th} component of the input $\boldsymbol{x}_j$ to the covariance kernel $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. These expressions can be extended to arbitrary derivatives simply by taking the relevant partial derivatives of $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. The SE covariance kernel given in equation (4) only depends on $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ through the quantity

$r = |\boldsymbol{\tau}| = |\boldsymbol{x}_i - \boldsymbol{x}_j|$; covariance kernels with this property are said to be *isotropic*. But, if $\boldsymbol{x}_i$ consists of position and time, for instance, we have no reason to expect the curve to behave the same in every direction. In general, one can make the substitution $r^2 \to \boldsymbol{\tau}^\mathsf{T}\mathbf{M}\boldsymbol{\tau}$ to introduce anisotropy into an isotropic covariance kernel, where $\mathbf{M}$ is any positive semidefinite matrix. This work takes $\mathbf{M}$ to be diagonal and uses the substitution

$$\frac{r^2}{\ell^2} \to \sum_{d=1}^{D} \frac{\tau_d^2}{\ell_d^2} = \sum_{d=1}^{D} \frac{(x_{i_d} - x_{j_d})^2}{\ell_d^2}, \tag{A.3}$$

where each dimension has its own covariance length scale ℓ_d.

Substituting equation (A.3), the general form of the SE covariance kernel is

$$k_{\mathrm{SE}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = k_{\mathrm{SE}}(\boldsymbol{\tau}) = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{d=1}^{D} \frac{\tau_d^2}{\ell_d^2}\right) \tag{A.4}$$

$$= \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{d=1}^{D} \frac{(x_{i_d} - x_{j_d})^2}{\ell_d^2}\right). \tag{A.5}$$

Turning the sum in the middle expression in equation (A.5) into a product yields

$$k_{\mathrm{SE}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma_f^2 \prod_{d=1}^{D} \exp\left(-\frac{\tau_d^2}{2\ell_d^2}\right). \tag{A.6}$$

Recall the Rodrigues representation of the (physicists') Hermite polynomials $\mathrm{H}_n(x)$ [53]:

$$\mathrm{H}_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2}, \tag{A.7}$$

where [54, 55]:

$$\mathrm{H}_n(x) = n! \sum_{l=0}^{\lfloor n/2 \rfloor} \frac{(-1)^l (2x)^{n-2l}}{l!(n-2l)!}. \tag{A.8}$$

Rearranging,

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} = (-1)^n \mathrm{H}_n(x) \mathrm{e}^{-x^2}. \tag{A.9}$$

To find the derivatives of equation (A.6), first find

$$\frac{\mathrm{d}^n}{\mathrm{d}\tau_d^n} \exp\left(-\frac{\tau_d^2}{2\ell_d^2}\right) = \left(\frac{1}{\sqrt{2}\ell_d}\right)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} = \left(\frac{-1}{\sqrt{2}\ell_d}\right)^n \mathrm{H}_n\left(\frac{\tau_d}{\sqrt{2}\ell_d}\right) \exp\left(-\frac{\tau_d^2}{2\ell_d^2}\right), \tag{A.10}$$

where the substitution $x = \tau_d/(\sqrt{2}\ell_d)$ was used. The derivative of equation (A.6) to arbitrary orders for each dimension is then

$$\frac{\partial^n k_{\mathrm{SE}}(\boldsymbol{\tau})}{\prod_{d=1}^{D} \partial \tau_d^{n_d}} = \sigma_f^2 \prod_{d=1}^{D} \left(\frac{-1}{\sqrt{2}\ell_d}\right)^{n_d} \mathrm{H}_{n_d}\left(\frac{\tau_d}{\sqrt{2}\ell_d}\right) \exp\left(-\frac{\tau_d^2}{2\ell_d^2}\right) \tag{A.11}$$

$$= k_{\mathrm{SE}}(\boldsymbol{\tau}) \prod_{d=1}^{D} \left(\frac{-1}{\sqrt{2}\ell_d}\right)^{n_d} \mathrm{H}_{n_d}\left(\frac{\tau_d}{\sqrt{2}\ell_d}\right), \tag{A.12}$$

where $n = \sum_{d=1}^{D} n_d$ and n_d is the order of derivative with respect to dimension d. What is needed to apply equations (A.1) and (A.2), however, are the derivatives with respect to x_{i_d} and x_{j_d}. Note that

$$\tau_d = x_{i_d} - x_{j_d}, \quad \frac{\partial \tau_d}{\partial x_{i_d}} = 1, \quad \frac{\partial \tau_d}{\partial x_{j_d}} = -1 \tag{A.13}$$

$$\frac{\partial^{n_i+n_j} \tau_d}{\partial x_{i_d}^{n_i} \partial x_{j_d}^{n_j}} = 0, \text{ for } n_i + n_j \geq 2. \tag{A.14}$$

Applying the chain rule to equation (A.12) then gives

$$\frac{\partial^n k_{\mathrm{SE}}(\boldsymbol{\tau}(\boldsymbol{x}_i, \boldsymbol{x}_j))}{\prod_{d=1}^{D} \partial x_{i_d}^{n_{i_d}} \prod_{d=1}^{D} \partial x_{j_d}^{n_{j_d}}} \tag{A.15}$$

$$= \frac{\partial^n k_{\mathrm{SE}}(\boldsymbol{\tau})}{\prod_{d=1}^{D} \partial \tau_d^{n_{i_d}+n_{j_d}}} \prod_{d=1}^{D} \left(\frac{\partial \tau_d}{\partial x_{i_d}} \right)^{n_{i_d}} \prod_{d=1}^{D} \left(\frac{\partial \tau_d}{\partial x_{j_d}} \right)^{n_{j_d}} \tag{A.16}$$

$$= k_{\mathrm{SE}}(\boldsymbol{\tau}) \prod_{d=1}^{D} \left(\frac{-1}{\sqrt{2}\ell_d} \right)^{n_{i_d}+n_{j_d}} \mathrm{H}_{n_{i_d}+n_{j_d}} \left(\frac{\tau_d}{\sqrt{2}\ell_d} \right) \cdot \prod_{d=1}^{D} (-1)^{n_{j_d}} \tag{A.17}$$

$$= k_{\mathrm{SE}}(\boldsymbol{\tau}) \prod_{d=1}^{D} \frac{(-1)^{n_{i_d}+2n_{j_d}}}{(\sqrt{2}\ell_d)^{n_{i_d}+n_{j_d}}} \mathrm{H}_{n_{i_d}+n_{j_d}} \left(\frac{\tau_d}{\sqrt{2}\ell_d} \right), \tag{A.18}$$

where $n = \sum_{d=1}^{D} n_{i_d} + \sum_{d=1}^{D} n_{j_d}$, n_{i_d} is the order of derivative with respect to dimension d of $\boldsymbol{x}_i$ and n_{j_d} is the order of derivative with respect to dimension d of $\boldsymbol{x}_j$. Equation (A.18) enables the computation of derivatives of *arbitrary* order for data of *arbitrary* dimensionality. A plot of the 1D SE covariance kernel and its first few derivatives is provided in figure A1.

References

[1] A. Bortolon et al. "Observation of Spontaneous Toroidal Rotation Inversion in Ohmically Heated Tokamak Plasmas". In: *Physical Review Letters* 97.23 (Dec. 2006), 235003. DOI: 10.1103/PhysRevLett.97.235003. URL: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.97.235003.

[2] J. E. Rice et al. "Observations of core toroidal rotation reversals in Alcator C-Mod ohmic L-mode plasmas". In: *Nuclear Fusion* 51.8 (2011), 083005. DOI: 10.1088/0029-5515/51/8/083005. URL: http://iopscience.iop.org/0029-5515/51/8/083005.

[3] J. E. Rice et al. "Ohmic energy confinement saturation and core toroidal rotation reversal in Alcator C-Mod plasmas". In: *Physics of Plasmas* 19.5 (2012), 056106. DOI: 10.1063/1.3695213. URL: http://aip.scitation.org/doi/abs/10.1063/1.3695213.

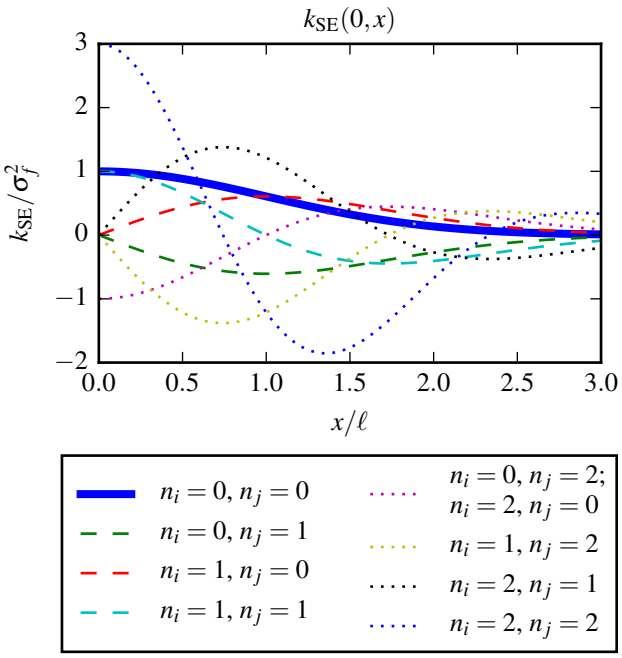

Figure A1: One-dimensional squared exponential covariance kernel and a few of its derivatives. The horizontal scale has been scaled by ℓ and the vertical scale has been scaled by σ_f^2 to remove the dependence on the hyperparameters. The solid blue curve gives $k_{\mathrm{SE}}(0,x)$ itself. The long-dashed curves give the functions necessary for dealing with first derivatives. The green dashed curve is $\partial k_{\mathrm{SE}}/\partial x_j = \mathrm{cov}[y(0), y'(x)]$. The red dashed curve is $\partial k_{\mathrm{SE}}/\partial x_i = \mathrm{cov}[y'(0), y(x)]$. The teal dashed curve is $\partial^2 k_{\mathrm{SE}}/\partial x_i\,\partial x_j = \mathrm{cov}[y'(0), y'(x)]$. The dotted curves give the functions necessary for dealing with second derivatives. The magenta dotted curve is $\partial^2 k_{\mathrm{SE}}/\partial x_i^2 = \partial^2 k_{\mathrm{SE}}/\partial x_j^2 = \mathrm{cov}[y''(0), y(x)] = \mathrm{cov}[y(0), y''(x)]$. The yellow dotted curve is $\partial^3 k_{\mathrm{SE}}/\partial x_i\,\partial x_j^2 = \mathrm{cov}[y'(0), y''(x)]$. The black dotted curve is $\partial^3 k_{\mathrm{SE}}/\partial x_i^2\,\partial x_j = \mathrm{cov}[y''(0), y'(x)]$. The blue dotted curve is $\partial^4 k_{\mathrm{SE}}/\partial x_i^2\,\partial x_j^2 = \mathrm{cov}[y''(0), y''(x)]$.

[4] J. E. Rice et al. "Non-local heat transport, rotation reversals and up/down impurity density asymmetries in Alcator C-Mod ohmic L-mode plasmas". In: *Nuclear Fusion* 53.3 (2013), 033004. DOI: 10.1088/0029-5515/53/3/033004. URL: http://iopscience.iop.org/0029-5515/53/3/033004.

[5] A. E. White et al. "Multi-channel transport experiments at Alcator C-Mod and comparison with gyrokinetic simulations". In: *Physics of Plasmas* 20.5 (2013), 056106. DOI: 10.1063/1.4803089. URL: http://scitation.aip.org/content/aip/journal/pop/20/5/10.1063/1.4803089.

[6] K. Ida et al. "Spontaneous toroidal rotation driven by the off-diagonal term of momentum and heat transport in the plasma with the ion internal transport barrier in LHD". In: *Nuclear Fusion* 50.6 (2010), 064007. DOI: 10.1088/0029-5515/50/6/064007. URL: http://iopscience.iop.org/0029-5515/50/6/064007.

[7] Felix I. Parra, Michael Barnes, and Peter J. Catto. "Sources of intrinsic rotation in the low-flow ordering". In: *Nuclear Fusion* 51.11 (2011), 113001. DOI: 10.1088/0029-5515/51/11/113001. URL: http://iopscience.iop.org/0029-5515/51/11/113001.

[8] Felix I. Parra, Michael Barnes, and Arthur G. Peeters. "Up-down symmetry of the turbulent transport of toroidal angular momentum in tokamaks". In: *Physics of Plasmas* 18.6 (2011), 062501. DOI: 10.1063/1.3586332. URL: http://scitation.aip.org/content/aip/journal/pop/18/6/10.1063/1.3586332.

[9] Y. Camenen et al. "Consequences of profile shearing on toroidal momentum transport". In: *Nuclear Fusion* 51.7 (2011), 073039. DOI: 10.1088/0029-5515/51/7/073039. URL: http://iopscience.iop.org/0029-5515/51/7/073039.

[10] Felix I. Parra et al. "Intrinsic rotation with gyrokinetic models". In: *Physics of Plasmas* 19.5 (2012), 056116. DOI: 10.1063/1.3699186. URL: http://scitation.aip.org/content/aip/journal/pop/19/5/10.1063/1.3699186.

[11] F. I. Parra et al. "Scaling of Spontaneous Rotation with Temperature and Plasma Current in Tokamaks". In: *Physical Review Letters* 108.9 (Mar. 2012), 095001. DOI: 10.1103/PhysRevLett.108.095001. URL: http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.108.095001.

[12] M. Barnes et al. "Intrinsic Rotation Driven by Non-Maxwellian Equilibria in Tokamak Plasmas". In: *Physical Review Letters* 111.5 (Aug. 2013), 055005. DOI: 10.1103/PhysRevLett.111.055005. URL: http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.111.055005.

[13] Michael Barnes. *Dependence of intrinsic rotation on higher-order profile gradients.* Personal communication. Aug. 2014.

[14] Jungpyo Lee, Felix I. Parra, and Michael Barnes. "Turbulent momentum pinch of diamagnetic flows in a tokamak". In: *Nuclear Fusion* 54.2 (2014), 022002. DOI: 10.1088/0029-5515/54/2/022002. URL: http://iopscience.iop.org/0029-5515/54/2/022002.

[15] J. P. Lee et al. "The effect of diamagnetic flows on turbulent driven ion toroidal rotation". In: *Physics of Plasmas* 21.5 (2014), 056106. DOI: 10.1063/1.4872322. URL: http://scitation.aip.org/content/aip/journal/pop/21/5/10.1063/1.4872322.

[16] Jungpyo Lee et al. "Turbulent momentum transport due to neoclassical flows". In: *Plasma Physics and Controlled Fusion* 57.12 (2015), 125006. DOI: 10.1088/0741-3335/57/12/125006. URL: http://iopscience.iop.org/article/10.1088/0741-3335/57/12/125006/.

[17] Felix I. Parra and Michael Barnes. "Intrinsic rotation in tokamaks: theory". In: *Plasma Physics and Controlled Fusion* 57.4 (2015), 045002. DOI: 10.1088/0741-3335/57/4/045002. URL: https://iopscience.iop.org/article/10.1088/0741-3335/57/4/045002/.

[18] J. C. Hillesheim et al. "Dependence of intrinsic rotation reversals on collisionality in MAST". In: *Nuclear Fusion* 55.3 (2015), 032003. DOI: 10.1088/0029-5515/55/3/032003. URL: http://iopscience.iop.org/0029-5515/55/3/032003.

[19] W. Dorland et al. "Electron Temperature Gradient Turbulence". In: *Physical Review Letters* 85.26 (Dec. 2000), pages 5579–5582. DOI: 10.1103/PhysRevLett.85.5579. URL: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.85.5579.

[20] A. M. Dimits et al. "Comparisons and physics basis of tokamak transport models and turbulence simulations". In: *Physics of Plasmas* 7.3 (Mar. 2000), pages 969–983. DOI: 10.1063/1.873896. URL: http://aip.scitation.org/doi/abs/10.1063/1.873896.

[21] *ITG: Cyclone base case*. URL: http://gs2.sourceforge.net/PMP/itg.html.

[22] M. A. Chilenski et al. "Improved profile fitting and quantification of uncertainty in experimental measurements of impurity transport coefficients using Gaussian process regression". In: *Nuclear Fusion* 55.2 (2015), 023012. DOI: 10.1088/0029-5515/55/2/023012. URL: http://iopscience.iop.org/0029-5515/55/2/023012.

[23] Mark Alan Chilenski. "Experimental Data Analysis Techniques for Validation of Tokamak Impurity Transport Simulations". PhD thesis. Massachusetts Institute of Technology, Oct. 2016.

[24] A. O'Hagan. "Curve Fitting and Optimal Design for Prediction". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 40.1 (1978), pages 1–42. URL: http://www.jstor.org/stable/2984861.

[25] David J. C. MacKay. "Introduction to Gaussian Processes". In: *Neural Networks and Machine Learning*. Edited by C. M. Bishop. NATO ASI Series. Kluwer Academic Press, 1998, pages 133–166. URL: http://www.inference.phy.cam.ac.uk/mackay/gpB.pdf.

[26] Michael L. Stein. *Interpolation of Spatial Data*. Springer Series in Statistics. New York, NY: Springer, 1999. ISBN: 978-1-4612-7166-6. DOI: 10.1007/978-1-4612-1494-6. URL: http://link.springer.com/book/10.1007/978-1-4612-1494-6.

[27] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006. ISBN: 978-0262182539. URL: http://www.gaussianprocess.org/gpml/.

[28] William H. Press et al. *Numerical Recipes: The Art of Scientific Computing*. 3rd edition. Cambridge: Cambridge University Press, 2007. ISBN: 978-0521880688. URL: http://numerical.recipes/.

[29] David Barber. *Bayesian Reasoning and Machine Learning*. The Edinburgh Building, Cambridge CB2 8RU, UK: Cambridge University Press, 2012. ISBN: 978-0521518147. URL: http://www.cs.ucl.ac.uk/staff/d.barber/brml/.

[30] Kevin P. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012. ISBN: 978-0262018029. URL: https://mitpress.mit.edu/books/machine-learning-0.

[31] Andrew Gelman et al. *Bayesian Data Analysis*. 3rd edition. Texts in Statistical Science. Boca Raton, FL: CRC Press, 2014. ISBN: 978-1439840955. URL: http://www.stat.columbia.edu/~gelman/book/.

[32] G. T. von Nessi et al. "Evidence cross-validation and Bayesian inference of MAST plasma equilibria". In: *Physics of Plasmas* 19.1 (2012), 012506. DOI: 10.1063/1.3677362. URL: http://scitation.aip.org/content/aip/journal/pop/19/1/10.1063/1.3677362.

[33] Dong Li et al. "Bayesian soft X-ray tomography using non-stationary Gaussian Processes". In: *Review of Scientific Instruments* 84.8 (2013), 083506. DOI: 10.1063/1.4817591. URL: http://scitation.aip.org/content/aip/journal/rsi/84/8/10.1063/1.4817591.

[34] J. A. Romero and J. Svensson. "Optimization of out-vessel magnetic diagnostics for plasma boundary reconstruction in tokamaks". In: *Nuclear Fusion* 53.3 (2013), 033009. DOI: 10.1088/0029-5515/53/3/033009. URL: http://iopscience.iop.org/0029-5515/53/3/033009.

[35] G. T. von Nessi, M. J. Hole, and the MAST Team. "A unified method for inference of tokamak equilibria and validation of force-balance models based on Bayesian analysis". In: *Journal of Physics A: Mathematical and Theoretical* 46.18 (2013), 185501. DOI: 10.1088/1751-8113/46/18/185501. URL: http://iopscience.iop.org/1751-8121/46/18/185501.

[36] G. T. von Nessi and M. J. Hole. "Using Bayesian analysis and Gaussian processes to infer electron temperature and density profiles on the Mega-Ampere Spherical Tokamak experiment". In: *Review of Scientific Instruments* 84.6 (2013), 063505. DOI: 10.1063/1.4811378. URL: http://scitation.aip.org/content/aip/journal/rsi/84/6/10.1063/1.4811378.

[37] A. Langenberg et al. "Forward Modeling of a High Resolution X-ray Imaging Crystal Spectrometer for the Wendelstein 7-X Stellarator". In: *41st European Physical Society Conference on Plasma Physics*. (Berlin, Germany, June 23–27, 2014). Edited by S. Ratynskaia et al. Volume 38F. European Conference Abstracts. European Physical Society, 2014, P1.074. URL: http://ocs.ciemat.es/EPS2014PAP/pdf/P1.074.pdf.

[38] Sehyun Kwak et al. "Bayesian modelling of the emission spectrum of the Joint European Torus Lithium Beam Emission Spectroscopy system". In: *Review of Scientific Instruments* 87.2 (Oct. 2016), 023501. DOI: 10.1063/1.4940925. URL: http://scitation.aip.org/content/aip/journal/rsi/87/2/10.1063/1.4940925.

[39] M. A. Chilenski. *gptools: Gaussian processes with arbitrary derivative constraints and predictions (online documentation)*. 2016. URL: http://gptools.readthedocs.org/.

[40] M. A. Chilenski. *gptools: Gaussian processes with arbitrary derivative constraints and predictions*. 2016. URL: https://github.com/markchil/gptools.

[41] M. A. Chilenski. *profiletools: Classes for working with profile data of arbitrary dimension (online documentation)*. 2016. URL: http://profiletools.readthedocs.org/.

[42] M. A. Chilenski. *profiletools: Classes for working with profile data of arbitrary dimension*. 2016. URL: https://github.com/markchil/profiletools.

[43] Nicholas Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (June 1953), pages 1087–1092. DOI: 10.1063/1.1699114. URL: http://scitation.aip.org/content/aip/journal/jcp/21/6/10.1063/1.1699114.

[44] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (1970), pages 97–109. DOI: 10.2307/2334940. URL: http://www.jstor.org/stable/2334940.

[45] M. A. Chilenski, I. C. Faust, and J. R. Walk. "eqtools: Modular, Extensible, Open-Source, Cross-Machine Python Tools for Working with Magnetic Equilibria". In: *Computer Physics Communications (in press)* (2016). DOI: 10.1016/j.cpc.2016.09.011. URL: http://www.sciencedirect.com/science/article/pii/S001046551630282X.

[46] M. A. Chilenski, I. C. Faust, and J. R. Walk. *eqtools: Tools for interacting with magnetic equilibria*. 2016. URL: http://eqtools.readthedocs.org/.

[47] M. A. Chilenski, I. C. Faust, and J. R. Walk. *eqtools: Python tools for magnetic equilibria in tokamak plasmas*. 2016. URL: https://github.com/PSFCPlasmaTools/eqtools.

[48] Jonathan Goodman and Jonathan Weare. "Ensemble samplers with affine invariance". In: *Communications in Applied Mathematics and Computational Science* 5.1 (2010), pages 65–80. DOI: 10.2140/camcos.2010.5.65. URL: http://msp.org/camcos/2010/5-1/p04.xhtml.

[49] Daniel Foreman-Mackey et al. "emcee: The MCMC Hammer". In: *Publications of the Astronomical Society of the Pacific* 125.925 (Mar. 2013), pages 306–312. DOI: 10.1086/670067. URL: http://www.jstor.org/stable/10.1086/670067.

[50] Neil H. Timm. *Applied Multivariate Analysis*. Springer Texts in Statistics. Springer, 2002. ISBN: 978-0-387-95347-2. DOI: 10.1007/b98963. URL: https://link.springer.com/book/10.1007/b98963.

[51] N. T. Howard et al. "Multi-scale gyrokinetic simulation of Alcator C-Mod tokamak discharges". In: *Physics of Plasmas* 21.3 (2014), 032308. DOI: 10.1063/1.4869078. URL: http://scitation.aip.org/content/aip/journal/pop/21/3/10.1063/1.4869078.

[52] N. T. Howard et al. "Validation of the gyrokinetic model in ITG and TEM dominated L-mode plasmas". In: *Nuclear Fusion* 53.12 (2013), 123011. DOI: 10.1088/0029-5515/53/12/123011. URL: http://iopscience.iop.org/0029-5515/53/12/123011.

[53] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists*. 6th edition. 30 Corporate Drive, Suite 400, Burlington, MA 01803, USA: Elsevier Academic Press, 2005. ISBN: 978-0123846549. URL: http://www.sciencedirect.com/science/book/9780123846549.

[54] *NIST Digital Library of Mathematical Functions*. URL: http://dlmf.nist.gov/.

[55] *NIST Digital Library of Mathematical Functions: 18.5 Explicit Representations*. URL: http://dlmf.nist.gov/18.5.