

# Bayesian Modelling of Microwave Radiometer Calibration on the example of the Wendelstein 7-X Electron Cyclotron Emission diagnostic

Udo Hoefel,<sup>1, a)</sup> Matthias Hirsch,<sup>1, b)</sup> Sehyun Kwak,<sup>1, 2</sup> Andrea Pavone,<sup>1</sup> Jakob Svensson,<sup>1</sup> Torsten Stange,<sup>1</sup> Hans-Jürgen Hartfuß,<sup>1</sup> Jonathan Schilling,<sup>1</sup> Gavin Weir,<sup>1</sup> Johan Willem Oosterbeek,<sup>1</sup> Sergey Bozhenkov,<sup>1</sup> Harald Braune,<sup>1</sup> Kai-Jakob Brunner,<sup>1</sup> Neha Chaudhary,<sup>1</sup> Hannes Damm,<sup>1</sup> Golo Fuchert,<sup>1</sup> Jens Knauer,<sup>1</sup> Heinrich Laqua,<sup>1</sup> Stefan Marsen,<sup>1</sup> Dmitry Moseev,<sup>1</sup> Ekkehard Pasch,<sup>1</sup> Evan R. Scott,<sup>1</sup> Fabian Wilde,<sup>1</sup> and Robert Wolf<sup>1, 3</sup> (the W7-X team)<sup>c)</sup>

<sup>1)</sup>Max Planck Institute for Plasma Physics, Wendelsteinstr. 1, D-17491 Greifswald, Germany

<sup>2)</sup>Department of Nuclear and Quantum Engineering, KAIST, Daejeon, 34141, Republic of Korea

<sup>3)</sup>Department of Astronomy and Astrophysics TUB, Hardenbergstr. 36, D-10623 Berlin, Germany

(Dated: 21 March 2019)

This paper reports about a novel approach to the absolute intensity calibration of an electron cyclotron emission (ECE) spectroscopy system. Typically, an ECE radiometer consists of tens of separated frequency channels corresponding to different plasma locations. An absolute calibration of the overall diagnostic including near plasma optics and transmission line is achieved with blackbody sources at LN<sub>2</sub> temperature and room temperature via a hot/cold calibration mirror unit. As the thermal emission of the calibration source is typically a few thousand times lower than the receiver noise temperature, coherent averaging over several hours is required to get a sufficient signal to noise ratio. A forward model suitable for any radiometer calibration using the hot/cold method and a periodic switch between them has been developed and used to extract the voltage difference between the hot and cold temperature source via Bayesian analysis. In contrast to the classical analysis which evaluates only the reference temperatures, the forward model takes into account intermediate effective temperatures caused by the finite beam width and thus uses all available data optimally. This allows the evaluation of weak channels where a classical analysis would not be feasible, is statistically rigorous and provides a measurement of the beam width. By using a variance scaling factor a model sensitive adaptation of the absolute uncertainties can be implemented, which will be used for the combined diagnostic Bayesian modelling analysis.

PACS numbers: 07.57.Kp

## I. INTRODUCTION

Microwave radiometers are applied in a wide range of research areas, ranging from atmospheric physics<sup>1</sup>, radio astronomy<sup>2</sup> to nuclear fusion research<sup>3,4</sup>. A common way to absolutely calibrate the radiometers is the hot/cold calibration method, which uses two reference temperatures to determine the calibration factors<sup>4</sup>, although often only a relative calibration with respect to a Michelson interferometer or a Thomson scattering diagnostic is done<sup>5</sup>. Due to the amount of components used in a radiometer it is challenging to get appropriate uncertainties for each component up to the data acquisition system (DAQ). Realistic uncertainties are especially important for modelling of multiple diagnostics, as too small or too big uncertainties will artificially shift the result.

In magnetically confined plasmas, electron cyclotron emission (ECE) is widely used to measure the electron temperature with high spatial and temporal resolution. The significance of the measurement is limited by the finite optical thickness of the plasma and black body radiation statistics.

At the optimized stellarator Wendelstein 7-X<sup>6,7</sup> (W7-X), the Bayesian combined diagnostic modelling is done within

the Minerva framework<sup>8</sup>. For the ECE system at W7-X, the hot/cold method mentioned above is used, with the whole calibration unit being located in the W7-X torus hall. A good estimate for the uncertainties of the ECE radiometer calibration is obtained by inverting the forward model of the calibration process described in this paper. The implementation allows easy automation of the whole calibration procedure given a calibration source making multiple reference temperatures available in the torus hall. A hot source could also be used, as planned for ITER<sup>9</sup>.

In plasma physics, it is typically the case that the measurements of multiple diagnostics depend, at least partially, on the same plasma parameters. To combine these measurements optimally, Bayesian analysis of forward models is the appropriate tool to use, as the classical, single diagnostic approach would neglect information shared between measurements of different diagnostics. A very nice example of an advanced Bayesian analysis approach can be found at ASDEX Upgrade<sup>10</sup>.

Sophisticated forward models of plasma ECE can be found for example at ASDEX Upgrade, see Rathgeber et al.<sup>11</sup> and Denk et al.<sup>12,13</sup>, as well as at JET, see Schmuck et al.<sup>14,15</sup>.

This paper is structured as follows: Section II introduces the ECE hardware setup at W7-X and the calibration procedure, section III explains the Minerva concepts and classes used to evaluate the obtained calibration raw data, section IV shows results from a hot/cold calibration, section III B gener-

<sup>a)</sup>Electronic mail: udo.hoefel@ipp.mpg.de

<sup>b)</sup>Electronic mail: matthias.hirsch@ipp.mpg.de

<sup>c)</sup>Members listed in R. C. Wolf et al. Nucl. Fusion 57, 102020 (2017).

alises the model such that it is in principle applicable to any radiometer calibrated with a hot/cold source and a rotating mirror and section V finally summarizes the results and discusses possible further enhancements.

## II. HARDWARE SYSTEM

The ECE system at W7-X<sup>16</sup> is measured at the outboard, low field side of the torus at a toroidal angle of  $\phi = 223^\circ$ . At a toroidal angle of  $\phi = 0^\circ$  the magnetic flux surfaces take on the shape of a bean, similar to the magnetic flux surfaces shown in figure 1, but symmetric around a horizontal cut at  $z = 0$ . Here, a sightline of the ECE truly perpendicular to the magnetic flux surfaces could be achieved. However, as the electron cyclotron resonance heating heats the plasma in these symmetry planes, the plasma center would not be accessible for ECE measurements. Thus, the ECE line of sight was chosen to have also a small toroidal component, while still near the bean shaped plane; due to the fivefold symmetry of W7-X,  $\phi = 216^\circ$  corresponds to  $\phi = 0^\circ$ . As a consequence, the ECE does measure a Doppler shift, and refraction plays a role for higher densities. However, the sightline has been optimized to reduce the influence of both these effects as much as possible<sup>17</sup>. A schematic representation of the whole setup is shown in figure 1, wherein the vacuum magnetic flux surfaces are shown for the standard magnetic configuration. A wideband optical Gauss telescope system with a  $1/e^2$  beam intensity width of approximately 20mm is used. The variation of the width over the sightline in the plasma ( $\pm 2$ mm) can be considered small. Exactly the same optical system including the vacuum windows is used below the W7-X experiment in the calibration unit, which is explained in more detail in section II A. The only difference between the two optical systems is the small tilt of the last plane in-vessel mirror to achieve a sightline perpendicular to the magnetic flux surfaces to suppress Doppler contributions. Instead of the inner plasma vessel, the calibration unit has its sightline leading to a steadily rotating, gold-coated mirror, which is surrounded by a highly microwave absorbent foam (ECCOSORB<sup>®</sup>) at room and liquid nitrogen temperature, see also figure 4. A detailed comparison of the in-vessel and calibration unit optics can be seen in figure 2 measured by performing a hot/cold calibration with each optical system setup in the lab with a short transmission line, and comparing the respective bit differences. Measurements on subsequent days yielded no drifts of the systematic differences. To clarify whether drifts on long timescales occur these measurements will be repeated in the future. A wire grid separates the incoming radiation in X and O mode, for which two separate but similar transmission lines exist. An oversized waveguide of approximately 23m length, including two tapers, 11 mitre-bends and one polarization tuner transmits the radiation from either the calibration unit or the plasma to the detection system outside the torus hall, allowing easy access during operation. The overall loss of this transmission line is 13.3 dB. The polarization tuner allows to adopt the mode to the radiometer input. Cross polarization coupling is on the order of 1 % to 2 %. A calibrated

noise source can be selected by a waveguide switch instead of the transmission line, which allows to calibrate with a higher signal to noise ratio at the expense of not taking the influence of components in front of the noise source switch into account. A Bragg reflection notch filter with at least 55 dB insertion loss within  $(140 \pm 0.5)$  GHz and approximately 5.3 dB insertion loss outside was used to block non-absorbed 140 GHz electron cyclotron resonance heating (ECRH) radiation<sup>18</sup>. Afterwards, the signal is down-converted to 4 GHz to 40 GHz via a phase locked loop (PLL) stabilized local oscillator (LO) at 122.06 GHz<sup>19</sup>. Via power dividers the power is split into 32 channels, which subsequently are band-pass filtered with center frequencies between 4.4 GHz and 39.6 GHz and a bandwidth of  $0.25 \text{ GHz} \leq \Delta f \leq 1.4 \text{ GHz}$ , chosen to adapt the radial resolution to the expected typical optical thickness of the plasma resulting in a resolution in real space between 0.5 cm and 1.5 cm<sup>17</sup>. After that, the signal passes through predetection amplifiers and the detection diodes. Highly linear post-detection amplifiers with a variable gain and an adjustable DC offset allow to choose a reasonable signal amplitude for each plasma discharge, therefore making maximum use of the range of the analog-to-digital converter (ADC). Subsequently, a low-pass filter with a 3 dB point at 550 kHz determines the maximum temporal resolution and ensures that no aliasing occurs. Finally, a 16 bit ADC with a sampling rate of up to 2 megasamples per second is used, before the data is stored in the central, immutable W7-X database by the data acquisition system (DAQ).

### A. Calibration Unit

Many ECE radiometers that are absolutely calibrated use a rotating blade to switch between two reference temperatures<sup>5,20</sup>. Another method to switch between the reference temperatures is given by a rotating mirror, as described by HartfuSS et al.<sup>4</sup>, and is sketched in figure 4. The advantage of these two methods in contrast to just recording data at one reference temperature for several minutes and then at another reference temperature for several minutes (both without using a chopper) is the decreased sensitivity to drifts of the measurement signal, which, if the drifts are not strictly linear, would change the ratio of the bit signal corresponding to the reference temperatures over time. The rotating mirror and the subsequently applied conditional averaging act as a bandpass which suppresses drifts on timescales larger than a rotation period, as the drifts correspond to a low frequency contribution. As drifts on timescales  $> 10$ s are not negligible at W7-X (see figure 3) despite having the electronics in a temperature controlled rack, the rotating mirror method has been chosen for W7-X as the temperature control allows only for temperature stability on the order of 1 K to 2 K. The advantage of the rotating mirror over a rotating chopper lies in the better symmetry of the intermediate temperatures that are measured when radiation from multiple radiation sources at different temperatures is collected. It should be noted that drifts on the magnitude observed here pose a considerable problem for long term plasma operation – either further measures to suppress drifts have to

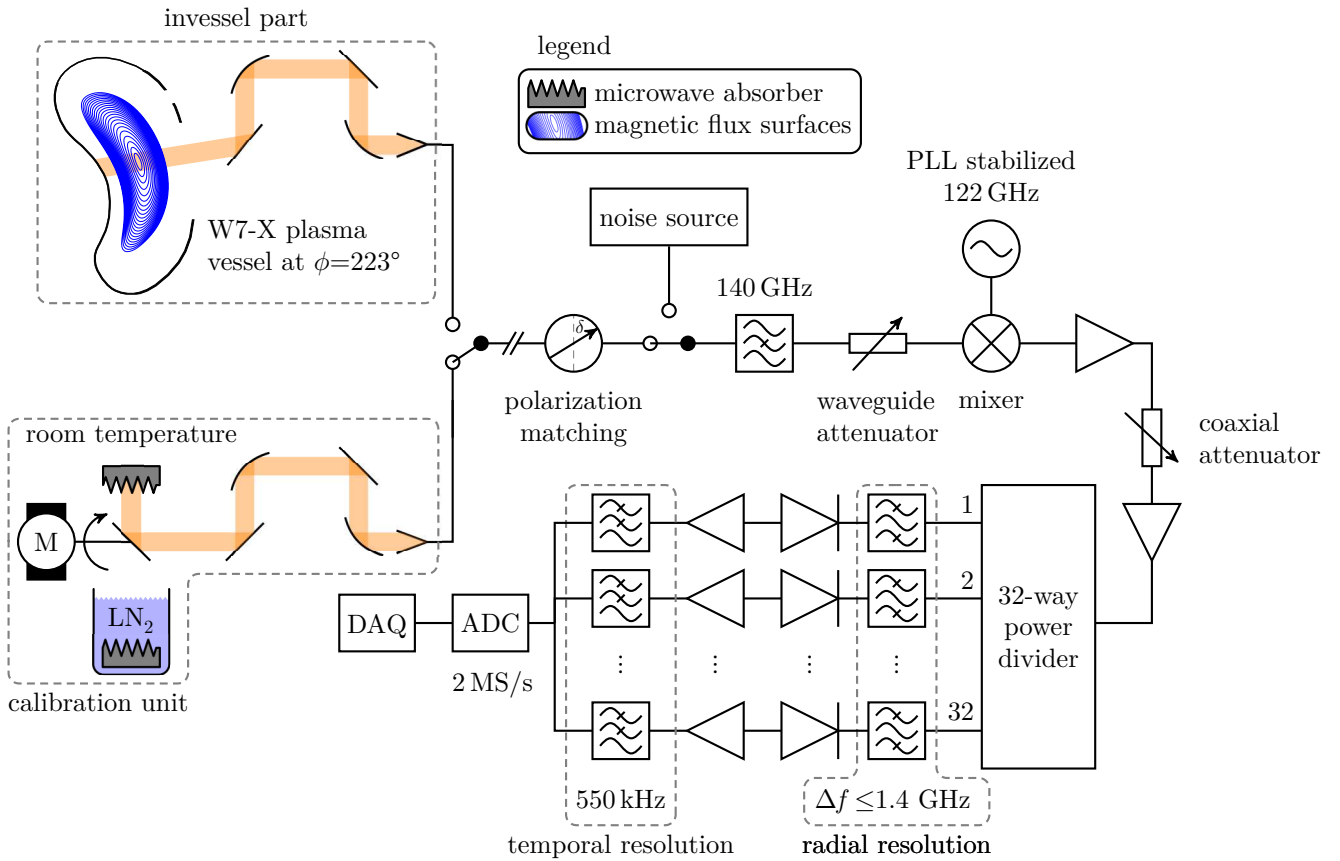


FIG. 1. Schematic drawing of the ECE diagnostic system at W7-X. The black lines around the shown magnetic flux surfaces indicate the vessel cross section. For a detailed description see section II.

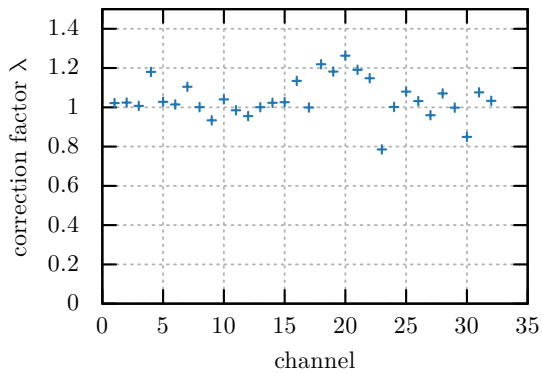


FIG. 2. Factors and estimated uncertainties required to match the rough estimated sensitivities from a hot/cold source calibration. This measurement has been obtained using the in-vessel optic in a lab setup, measuring each channel separately. The origin of the deviations is currently not understood. Tests do not indicate an origin caused by standing waves.

be taken, or a regularly repeated offset determination within a discharge has to be performed. For W7-X, it is planned to repeatedly close the shutter in the planned 30 min plasma discharges to correct the offset.

The calibration unit of W7-X contains a gold-coated brass

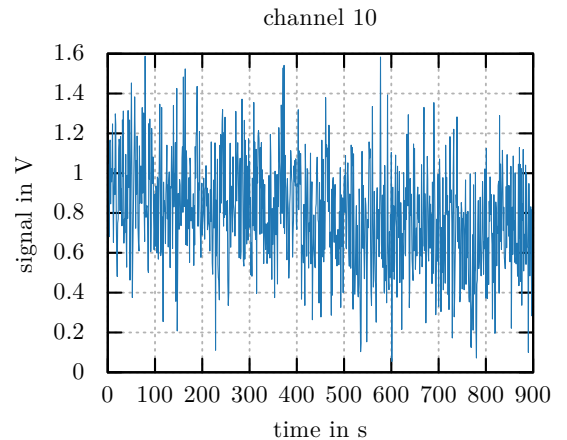


FIG. 3. Example of the drifts of an ECE channel during a few minutes. One can see a notable drift of the average signal, thus measuring hot and cold temperatures separately for several minutes would yield considerably larger uncertainties if the drifts are not strictly linear. Channel 10 has a signal difference between the two reference temperatures of around 4.7 mV. Only every 1000<sup>th</sup> point is shown.

mirror rotating with approximately 3.6 Hz, see figure 4. Arranged cylindrically around the mirror, a microwave absorber guarantees a black body emitter at room temperature  $T_{RT}$ ,

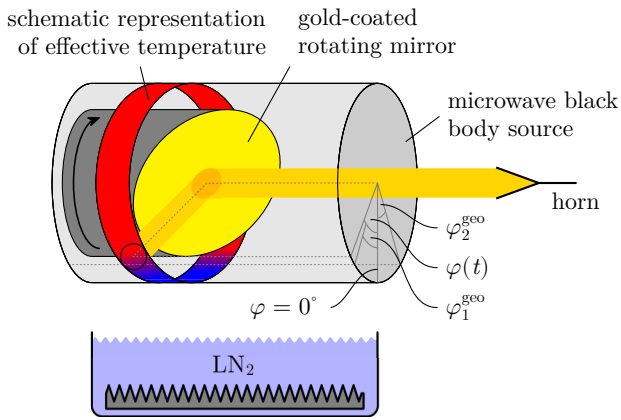


FIG. 4. Schematic drawing of the ECE calibration unit. Not shown is the optical system between the rotating mirror part of the calibration unit and the microwave antenna, which is identical to the invessel optical system. The path that is “seen” by the microwave antenna is shown in light orange. The microwave antenna is characterized by a Gaussian beam. The beam is reflected at the gold-coated mirror and finally “sees” either room or liquid nitrogen temperature, the latter being produced by a liquid nitrogen tank underneath the rotating mirror. Due to the finite size of the beam, the effectively measured temperatures are smeared out at the hot/cold edges, as different parts of the beam “see” different temperatures.

which is kept at  $(294.45 \pm 3.5)$  K in the torus hall. However, a small part at the lower side of this cylinder is cut to allow the observation of a stainless steel container thermally insulated by styrofoam. The inner wall of the stainless steel container is lined with a microwave absorber. The cold reference temperature is not directly given by the temperature of the liquid nitrogen,  $T_{\text{LN}_2} = (77.2 \pm 0.5)$  K, as water vapour (assumed to be at  $T_{\text{H}_2\text{O}} = (280 \pm 10)$  K with an uniform emissivity of  $0.01 < \epsilon_{\text{H}_2\text{O}} < 0.03$ ) accumulates above the liquid nitrogen reservoir. Moreover, the temperature of the mirror needs to be taken into account (corresponding to room temperature described above, with an emissivity  $0.01 < \epsilon_{\text{mirror}} < 0.03$ ). This leads to an effective temperature difference between the hot and the cold source of about 205 K that is used in further calculations. Details in the effective temperature estimation are given in section III B.

### III. CALIBRATION MODELLING

To obtain the physical quantities that are of interest for the evaluation of an ECE calibration, namely the calibration factor of each individual channel (that is the inverse sensitivity) and the effective beam width, the calibration procedure itself is modeled. To evaluate multiple ECE channels in a consistent way, a forward model predicting the different channel sensitivities has to be used. Using Bayes’ formula<sup>21</sup> the sensitivi-

ties and the effective beam widths can be inferred by

$$P(F|D) = \frac{\overbrace{P(D|F)}^{\text{likelihood}} \overbrace{P(F)}^{\text{prior}}}{\underbrace{P(D)}^{\text{evidence}}}, \quad (1)$$

wherein  $P(\cdot)$  represents a probability,  $F$  the free parameters to be inferred and  $D$  the data. A priori knowledge about the free parameters is encoded in the *prior* probability distributions. The probability of the measured data, given the free parameters, is called the *likelihood*. The normalization factor in the denominator, often called *evidence*, is important for model comparison. The *posterior* is the term we want to calculate, as it yields the probability distribution of the sought-after free parameters given the data. As the evidence does not change the shape of the posterior, one can neglect that term if one is not interested in comparing models explicitly, for which the different models would have to be normalized with a penalty on complexity (that is, by applying Occams razor).

#### A. Graphical Models

Graphical models are a powerful tool to describe the conditional dependency structure of a probabilistic model<sup>22</sup>. A Minerva graphical model is a Bayesian network, more precisely, a directed acyclic graph. It consists of nodes, which are connected via arrows, that are used to evaluate, for example, the calibration processes in dependence of their parent nodes. The nodes can be either deterministic (the rectangles in figure 6) or probabilistic (for example the uniform and normal prior probabilistic nodes shown in blue in figure 6). The edges specify explicitly the dependencies that a node has on other nodes. Together, the model encodes the whole joint probability of the model and the data. Minerva is a Java based general Bayesian modelling framework used among several other large scale experiments around the world, for example at the Joint European Torus<sup>23</sup> (JET) and the Mega-Ampere Spherical Tokamak<sup>24</sup> (MAST) and, as the main inference framework, at W7-X. The models in this paper are constructed within this framework, and also evaluated within it, as Minerva offers numerous different techniques for inversions, such as maximum a posteriori (MAP) and Markov Chain Monte Carlo (MCMC) methods. Do note that these models can be used to generate sample predictions that may be used for example for neural net training sets<sup>25</sup>.

A correction factor  $\alpha$  allows to scale the variance to obtain the most probable uncertainties for the prediction. As  $\alpha$  is a free parameter that changes the variance,  $\alpha$  is a hyperparameter, as different uncertainties correspond to different models. The explicit expression, for uncertainties following a Gaussian distribution, is given by

$$P(D|\mathbf{p}, \alpha) = \frac{1}{\sqrt{2\pi\alpha\sigma^2}} \exp\left(-\frac{(D-F(\mathbf{p}))^2}{2\alpha\sigma^2}\right), \quad (2)$$

with  $P(\cdot)$  the probability,  $D$  the observed data,  $\mathbf{p}$  other free parameters,  $\alpha$  the factor that scales the variance,  $\sigma^2$  the vari-

ance and  $F(\cdot)$  the forward function. However, using  $\alpha$  does not imply that no effort in increasing the understanding of uncertainties should be made.

For comparison of different modelling approaches, different descriptions of the calibration process were implemented. The following properties of the model have been combined, creating a total of eight structurally different models: i) Evaluation of individual channels respectively multiple channels simultaneously (indicated from here on with the keywords 'single' respectively 'multi'), ii) with and without the use of a channel specific scaling factor  $\alpha_i$  to scale the variance of the observed conditionally averaged signal and iii) by having a channel specific beam width or a beam width fixed for all channels by a scaling following

$$w(f) = w(140\text{GHz})\sqrt{140\text{GHz}/f}, \quad (3)$$

as expected from broadband Gauss telescope optics, indicated from here on with the keywords 'shared' respectively 'individual'. It should be noted that technically the models with a variance scaling factor  $\alpha_i$  describe a continuum of assumed uncertainty models.

The advantage of evaluating all channels simultaneously is the consistency gained for channel independent parameters, namely the beginning and ending of the hot temperature source, and, depending on the model, the beam width. The disadvantage that comes with evaluating multiple channels simultaneously lies in the curse of dimensionality – the evaluation time increases notably. Exemplarily, the graphical model for the multiple channel case with scaling factors  $\alpha_i$  and channel specific beam widths is shown in a simplified way in figure 6.

The single channel evaluation will be compared with the multi channel evaluation in section IV. The evaluation strategy in general is the following: first get reasonable estimates for the starting parameters via the datasource, then do a maximum a posteriori (MAP) inversion (for example using the Hooke and Jeeves pattern search algorithm<sup>26</sup>), of which the result in turn is used as a starting point for the MCMC inversion with a sufficient burn in and an adaptive Metropolis adapter<sup>27</sup> that is deactivated once the MCMC chain is stable. The burn in is a period at the beginning of an MCMC in which all samples are discarded from further analysis. It is not strictly necessary, but it helps in the interpretation of the result, as otherwise, if one starts in an unlikely position of the posterior distribution and does not run the chain long enough, the result will be biased towards these unlikely values. Do note that the MCMC in use is essentially a black box MCMC, thus the only way to ensure convergence instead of just pseudo-convergence is by using sufficiently long runs<sup>28</sup>. Therefore, all MCMCs have been run with a burn in of at least 1 million iterations, until the MCMC traces of the free parameters did not show notable drifts of the running average of the logarithm of the probability density function. All MCMCs ran so far show convergence after at most 800 000 MCMC iterations, the single channel evaluations typically after less than 10 000. This allows for a realistic estimation of the uncertainties of the drop of the measured signal associated with the decrease in radiation temperature as produced by the switch

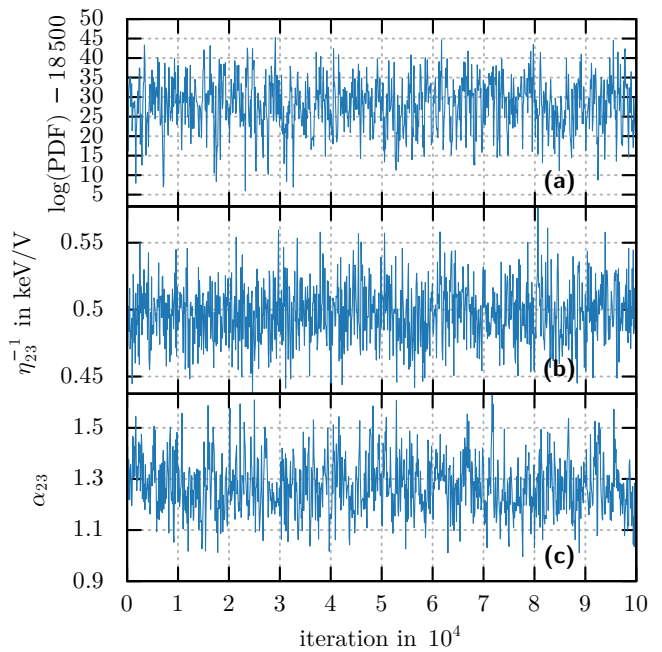


FIG. 5. (a) shows the logarithm of the probability density function of an MCMC chain of the full model with variance scaling factors. (b) and (c) show the MCMC chain of the calibration factor  $\eta_{23}^{-1}$  and the variance scaling factor  $\alpha_{23}$ . One can see that the values do not show jumps to a notably different phase space part, indicating that the chain converged. Only every 100<sup>th</sup> point is shown to keep the plot data volumes at a reasonable size.

from room temperature to the cold temperature. The logarithm of the probability density function, which is a measure for how probable a specific realisation is, of an MCMC chain of the full model, and example plots of each the calibration and scaling factor traces are shown in figure 5.

## B. Model for W7-X

To calculate the temperature difference between the room temperature,  $T_{RT}$ , and the effective cold temperature, one needs to take the influence of the water vapour emissivity,  $\epsilon_{H_2O}$ , and the mirror emissivity,  $\epsilon_{Mirror}$ , into account. The effective temperature after the radiation passes through the water vapour (at  $T_{H_2O}$ ) forming above the liquid nitrogen (at  $T_{LN_2}$ ) is calculated via

$$T_{eff}^{vapour} = T_{LN_2} + \epsilon_{H_2O}(T_{H_2O} - T_{LN_2}). \quad (4)$$

The effective cold temperature after the mirror is given by

$$T_{eff}^{cold} = T_{eff}^{vapour} + \epsilon_{mirror}(T_{eff}^{hot} - T_{eff}^{vapour}). \quad (5)$$

The emissivity of the mirror only adds to the effective cold temperature and not the effective hot temperature, as the mirror is already at the same temperature as the effective hot temperature. In case the hot reference temperature would be at a temperature different from the mirror temperature it would have to be taken into account there as well. Note that this

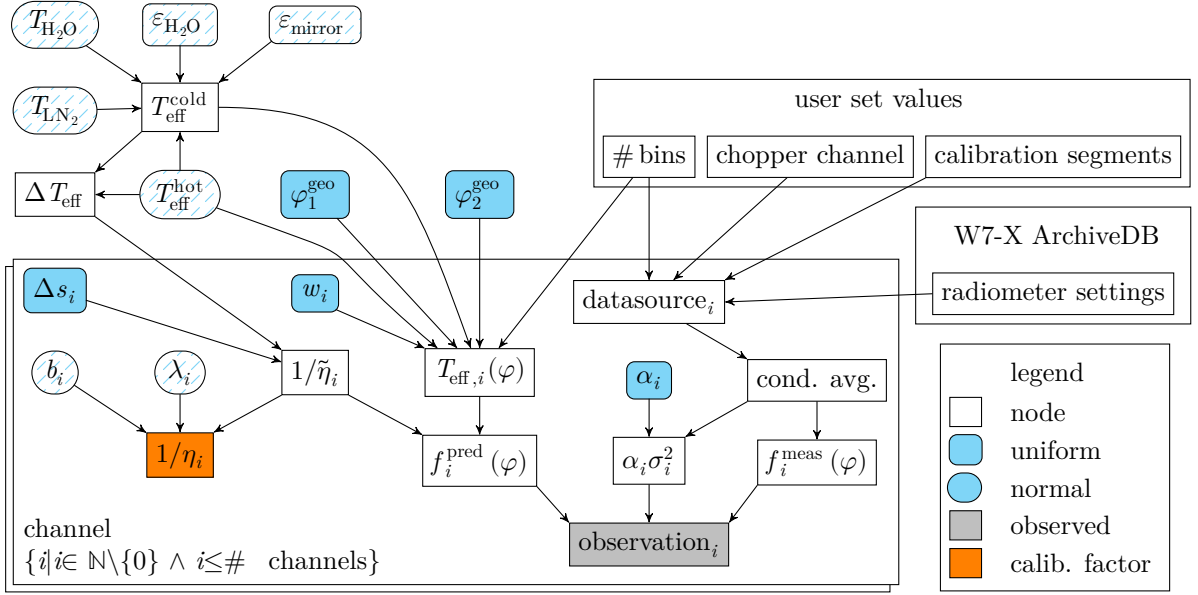


FIG. 6. The simplified Minerva graph showing the dependencies used in the model. Free parameters have a cyan background, while the shape specified in the legend corresponds to the used prior distribution. Do note that the nodes with a hatched background do sample completely randomly from their distribution, thus propagating the corresponding uncertainties without increasing dimensionality by parameters which do not influence the prediction. This approach was used for the temperatures and emissivities.

works because the signals are recorded in the frequency range where the Jeans law approximation holds. This allows the calculation of the channel specific calibration sensitivity (respectively, the sensitivity  $\eta_i$  scaled with the attenuation and postdetection amplification as chosen for this channel during the calibration)

$$\tilde{\eta}_i = \frac{\Delta s_i}{\Delta T_{\text{eff}}} = \frac{\Delta s_i}{T_{\text{eff}}^{\text{hot}} - T_{\text{eff}}^{\text{cold}}}, \quad (6)$$

wherein  $\Delta s_i$  represents the signal change in bits caused by the temperature difference  $\Delta T_{\text{eff}}$ . The index  $i$  denotes the channel specific parameters. Moreover, the channel specific expected effective temperature, depending on the mirror position, is given by

$$T_{\text{eff},i}(\varphi) = \frac{\int \hat{T}_{\text{eff}}(\hat{\varphi}) g_{\varphi}(w_i, \hat{\varphi}, \varphi) d\hat{\varphi}}{\int g_{\varphi}(w_i, \hat{\varphi}, \varphi) d\hat{\varphi}}, \quad (7)$$

where  $\hat{T}_{\text{eff}}$  is the effective temperature at a given mirror position without taking the finite width of the Gaussian beam, defined by the horn characteristic, into account,

$$\hat{T}_{\text{eff}}(\hat{\varphi}) = \begin{cases} T_{\text{eff}}^{\text{cold}} & \text{for } \hat{\varphi} < \varphi_1^{\text{geo}} \\ T_{\text{eff}}^{\text{hot}} & \text{for } \varphi_1^{\text{geo}} \leq \hat{\varphi} \leq \varphi_2^{\text{geo}} \\ T_{\text{eff}}^{\text{cold}} & \text{for } \hat{\varphi} > \varphi_2^{\text{geo}} \end{cases}, \quad (8)$$

and  $g_{\varphi}$  the weight of each  $\hat{T}_{\text{eff}}$  assuming a perfect Gaussian beam horn characteristic. The start of the hot source phase of the signal is described by  $\varphi_1^{\text{geo}}$ , the end by  $\varphi_2^{\text{geo}}$  (see figure 4). The beam width of the microwave antenna characteristic is denoted by  $w_i$ .

This in turn allows to predict the measured bit signal in dependence of the mirror angle,

$$f_i^{\text{pred}}(\varphi) = T_{\text{eff},i}(\varphi) \tilde{\eta}_i. \quad (9)$$

In practice, it can be useful to subtract the mean of the predicted signal to avoid dependencies on the effective temperatures where not necessary, see appendix. The observation consists of the conditionally averaged and binned signal  $f_i^{\text{meas}}(\varphi)$  in bit and the variance scaled by a factor  $\alpha_i$ .

The calibration factor is given by

$$\frac{1}{\eta_i} = \frac{\lambda_i b_i G_i}{\tilde{\eta}_i} 10^{-(R_{\text{RF}} + R_{\text{IF}})/10}. \quad (10)$$

This takes the following quantities into account: The measured differences between the invessel and calibration optics,  $\lambda_i$ , the measured bit to volt conversion factor,  $b_i$ , the post detection amplification chosen during calibration,  $G_i$ , the setting of the waveguide attenuator right in front of the radiometer in decibel,  $R_{\text{RF}}$ , and the setting of the attenuator at the intermediate frequency device in decibel,  $R_{\text{IF}}$ . For calibrating plasma measurements one needs to rescale this factor with the appropriate gains and attenuator settings used during the measurement. The reason for including this branch in the model is twofold: i) it allows direct extraction of the sought after quantity, without having to implement separate uncertainty propagation for  $\lambda_i$  and  $b_i$  and ii) preparing for future evaluations of multi diagnostic calibration factors, in which case one can simply extend this branch by supplying ECE raw plasma data to get an electron temperature profile that can also be supplied for example via Thomson scattering<sup>29</sup>. This would not be a simple cross calibration, but would rather combine the

diagnostic specific calibration models and their corresponding plasma forward models, thus taking all information optimally into account. The calibration factors obtained that way are inherently consistent within the frame of the model.

In the Bayesian formalism this leads to (for the model shown in figure 6)

$$P(\varphi^{\text{geo}}, \Delta \mathbf{s}, \mathbf{w}, \alpha | \tilde{\mathbf{D}}) = \frac{P(\tilde{\mathbf{D}} | \varphi^{\text{geo}}, \Delta \mathbf{s}, \mathbf{w}, \alpha) P(\varphi^{\text{geo}}, \Delta \mathbf{s}, \mathbf{w}, \alpha)}{P(\tilde{\mathbf{D}})}, \quad (11)$$

with

$$\varphi^{\text{geo}} = (\varphi_1^{\text{geo}}, \varphi_2^{\text{geo}}), \quad (12)$$

$$\Delta \mathbf{s} = (\Delta s_1, \dots, \Delta s_n), \quad (13)$$

$$\mathbf{w} = (w_1, \dots, w_n), \quad (14)$$

$$\alpha = (\alpha_1, \dots, \alpha_n), \quad (15)$$

$$\tilde{\mathbf{D}} = (\mathbf{D}_{\text{ECE},1}, \dots, \mathbf{D}_{\text{ECE},n}), \quad (16)$$

where  $n$  corresponds to the number of ECE channels,  $\varphi^{\text{geo}}$  to the angles at which the central line of sight switches from the hot source to the cold source and vice versa,  $\Delta \mathbf{s}$  to the channel specific change in the bit signal observed when switching the temperature sources,  $\mathbf{w}$  to the channel specific Gaussian beam width,  $\alpha$  to the channel specific variance scaling factors and  $\tilde{\mathbf{D}}$  to the channel specific conditionally averaged measured data. The conditional averaging is done in the following way: First, the average number of data in one mirror rotation is calculated. The time series is then split at each falling edge of the chopper signal. The resulting individual rotation measurements have their mean removed, are rescaled and sorted into the number of bins determined in the first step. Finally, dividing by the number of rotations yields the conditionally averaged data required for the analysis. Due to the small signal to noise ratio of the raw data, many rotations of the rotating mirror need to be acquired to extract enough information for meaningful results. Note that the other models differ, for example by not using the variance scaling factors.

### C. Generalised model

The model described previously uses some simplifications that can easily be dropped to generalise the model. For instance, one can drop the assumption that the problem is one dimensional and that there are only two reference temperatures. This allows easy extension to three (or more) reference temperatures for example by adding a hot ceramics hot source. Switching to cylindrical coordinates is a sensible approach for a geometry similar to the one presented here, thus introducing  $z$  along the horizontal axis of the cylinder shown in figure 4, and  $r$  as the radius. Assuming  $i \times j$  reference temperatures leads to a definition of the effective temperature  $\hat{T}_{\text{eff}}(\hat{\phi}, \hat{z}) = T_{\text{eff},ij}$ , where  $T_{\text{eff},ij}$  is the effective temperature valid for  $\varphi_i^{\text{geo}} \leq \hat{\phi} \leq \varphi_{i+1}^{\text{geo}}$  and  $z_j^{\text{geo}} \leq \hat{z} \leq z_{j+1}^{\text{geo}}$ . The second dimension is represented by  $\hat{z}$ . As in the model described above,  $T_{\text{eff},ij}$  can be the result of multiple layers contributing to the

effective temperature at the selected coordinates, such that a dependency on  $r$  might occur as well. The calibration sensitivity  $\tilde{\eta}(T)$  does not necessarily have to be linear, however, one will have to use free parameters for the temperatures and emissivities in this case, as the prediction will no longer be independent of these parameters. In general, any instrument function  $g$  can be used to calculate the appropriate weighted effective temperature that the radiometer would see by looking at  $(\phi, z)$ , therefore

$$T_{\text{eff}}(\phi, z) = \frac{\iint \hat{T}_{\text{eff}}(\hat{\phi}, \hat{z}) g(\hat{\phi}, \phi, \hat{z}, z) d\hat{\phi} d\hat{z}}{\iint g(\hat{\phi}, \phi, \hat{z}, z) d\hat{\phi} d\hat{z}}, \quad (17)$$

which in combination with the calibration sensitivity allows the calculation of the prediction.

### D. Limitations

Practical limitations are given for example by the number of bins as well as the number of channels. If the sampling rate notably exceeds 1 kHz the evaluation slows down considerably as well, as the data that has to be loaded for the preprocessing increases accordingly. A practical number of bins can be determined automatically, which gives a value close to the average number of data points per mirror rotation. As there are computationally expensive steps involved for each bin, increasing the number of bins also leads to an increase of required computation time. The full 98D model includes all 32 channels and the scaling of the variance, and takes roughly 230 hours with a Intel<sup>®</sup> Xeon<sup>®</sup> CPU E5-2660 v4 @ 2GHz on a virtualized linux server. By evaluating each channel independently the required CPU time is reduced to about 21 hours, such that with full parallelization the evaluation time can go down to around 40 minutes. This simplified model comes at the price of a generally smaller consistency and larger uncertainties for the weaker channels. However, the differences are negligible for reasonably strong channels (differences in the calibration factor are typically below 1%). For weaker channels the difference can reach about 10%. This is due to the stronger channels keeping the geometrical factors more or less fixed, such that the impact on channels where the geometrical information is more concealed in noise profit the most. The single channel evaluation routine provides a pragmatic approach to obtain calibration factors if time requirements prohibit the full model use.

## IV. RESULTS

A typical excerpt of the calibration timetrace for a sensitive channel can be seen in figure 7b, while the corresponding chopper signal is shown in figure 7a. There are three points to consider: i) the chopper signal does not correspond to the full width of the cold phase, the real hot/cold duty cycle is approximately 0.2, given by the calibration unit geometry, increasing that value further would require a significant modification of the calibration unit, ii) even for the most sensitive channel

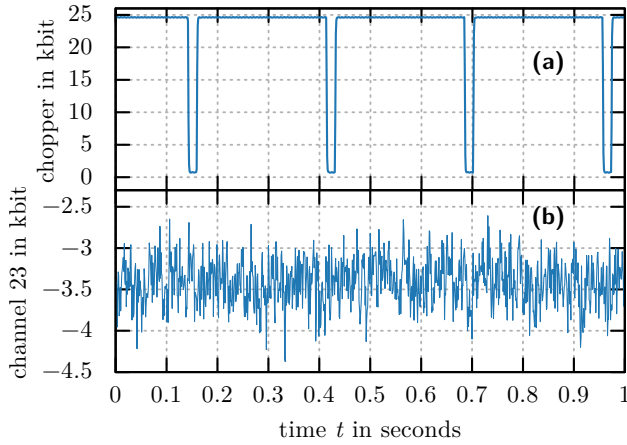


FIG. 7. **(a)** shows the chopper signal measured by a photo diode, **(b)** the raw signal of channel 23, which is a sensitive channel. The signal-to-noise ratio is typically of the order of 1/50. The background signal results from the broadband noise of the IF amplifiers right after the mixer, subsequently measured for each frequency bin by the detector diode.

the signal difference associated with the chopper channel cannot be seen directly, confirming that more elaborated analysis techniques are necessary and iii) no relevant drift within one rotation period can be observed. As said before, notable drifts were seen in some cases on timescales on the order of 10s (see figure 3). The conditional averaged signal in figure 8 supports the conclusion that there is no relevant drift within one period. The blue curve in figure 8a) corresponds to the measured and subsequently conditional averaged signal of the strong channel 23,  $f_{23}^{\text{meas}}(\varphi)$ , while the orange curves are samples from the graph that has been set to the mean values obtained from the previously run MCMC inversion, so these are the predictions,  $f_{23}^{\text{pred}}(\varphi)$ . Correspondingly, figure 8b) shows the weakly sensitive channel 11. It is important to note that these samples are calculated from the model that allowed the scaling of the prediction variance. Each orange point in figure 8 corresponds to a predicted effective temperature scaled by the calibration sensitivity, with the offset of a whole period being removed.

The measured bit values at each  $\varphi$  value are illustrated in figure 9, where the bit values have been scaled to represent voltages. It should be noted that the uncertainties on the voltage axis are scaled with the channel specific variance scaling factor  $\alpha$ . For the sake of clarity error bars are only shown in the intermediate temperature range. The orange curve shows the sensitivity with its uncertainties as calculated from the graphical model. Remaining deviations might be caused by 50Hz noise (or its higher harmonics) that are not completely notched out by the bandpass filter properties of the conditional average. The plot highlights the advantage of this analysis method: While no other radiometer calibration approach known to the authors uses the data that is taken when the horn pattern collects radiation from more than one reference temperature, this method allows to estimate the effective temperature (and corresponding uncertainties) reducing the overall uncertainty and predicting the frequency dependent beam

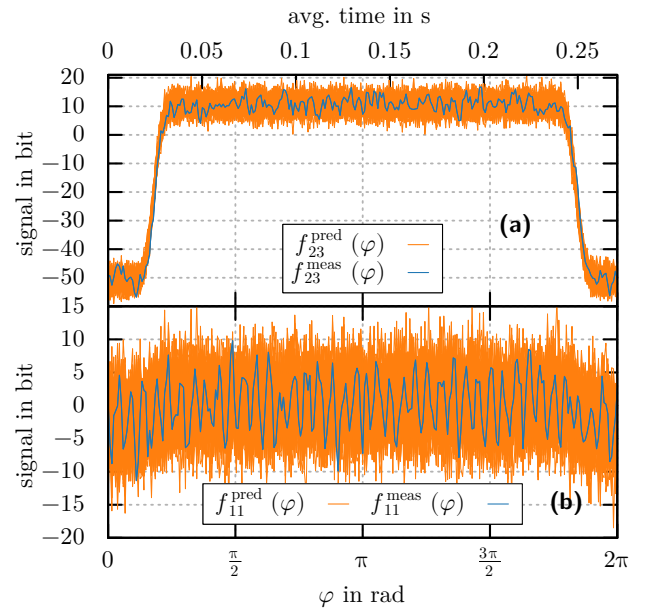


FIG. 8. **(a)** shows the sensitive channel 23, while **(b)** depicts the weakly sensitive channel 11.  $\varphi = 0$  corresponds to a mirror position “looking” at the center of the liquid nitrogen cold source. One can see the measured (blue) conditionally averaged and binned signal for both channels. The predictions (orange) are 100 iid Monte Carlo samples each, taken after the free parameters of the graph have been set to their mean values (i.e. a point estimate of the posterior predictive).

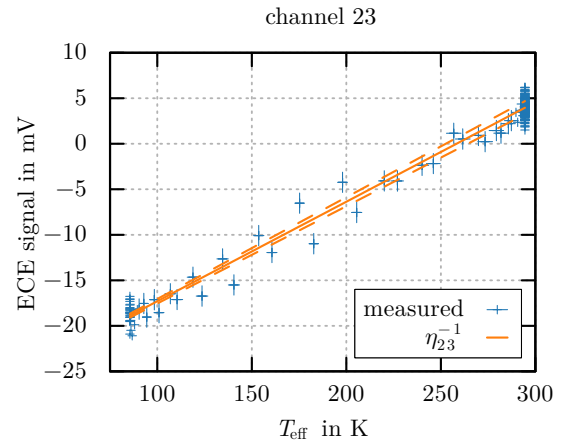


FIG. 9. The voltage values in dependence of the effective temperature. Do note that the plotted voltage uncertainties correspond to  $b_{23}\sqrt{\alpha_{23}\sigma_{23}^2}$  with  $b_{23}$  the bit to Volt conversion factor. The temperature uncertainties correspond to the uncertainties given by the priors.

width. However, one has to keep in mind that this is valid only as long as the Jeans law approximation is valid. For the radiometer and the reference temperatures used here this is a very good approximation.

The kernel density estimation<sup>30,31</sup> of the calibration factors given by the MCMC is shown in figure 10. One of the strengths of an MCMC based evaluation is that one can get



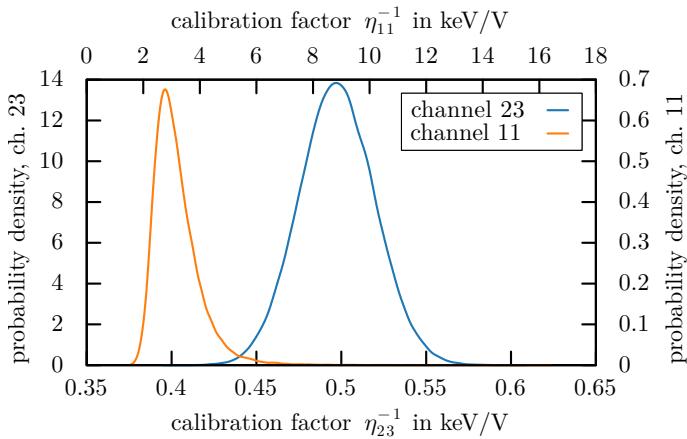


FIG. 10. The kernel density estimates of the MCMC samples of the calibration factors of the strong channel 23,  $\eta_{23}^{-1}$ , and the weak channel 11,  $\eta_{11}^{-1}$ . It is clearly visible that the distribution of the weak channel deviates from a normal distribution by having a pronounced tail towards larger calibration factors. In contrast, the stronger channel closely follows a normal distribution.

posterior distributions that are non-Gaussian as well. However, in the case shown here a Gaussian fit is a reasonable approximation to the posterior for the strong channel 23, while for channel 11 the posterior deviates notably from a Gaussian distribution. To quantify the deviation for channel 11: The mean of a Gaussian fit to the kernel density estimate is 3.03 keV/V, while the maximum of the kernel density estimate is at 2.92 keV/V (roughly 96.4% of the Gaussian mean). For the sake of simplicity the results of a Gaussian fit are used for all higher level analysis, although it would mark a gain in consistency, especially for the weak channels, to use the asymmetric uncertainties originating from the MCMC.

### A. Comparison of single and multi channel evaluation

As it is not ab initio clear how large the differences between the models of varying complexity are, and thus which model is appropriate for practice, a careful comparison is shown in this section.

Figure 11a shows the calibration factor, which is the inverse sensitivity, for each channel. One can see that the calibration factors vary over more than two orders of magnitude. A single mixer is used for the whole spectrum to allow for a better correlation analysis<sup>19</sup>. This is unusual as many ECE systems use multiple mixers to avoid frequencies above 18GHz after mixing<sup>32</sup>. The single mixer approach leads to intermediate frequencies up to 40GHz which need to be detected. The low sensitivities for higher frequencies might at least partially originate from different cables used for frequencies below 18GHz (corresponding to frequencies below 140GHz in the shown spectrum) and above 18GHz. The cable frequency response damps higher frequencies more. The conversion efficiency of the extreme broadband mixer also drops for frequencies above 140GHz. Individual diode sensitivities are

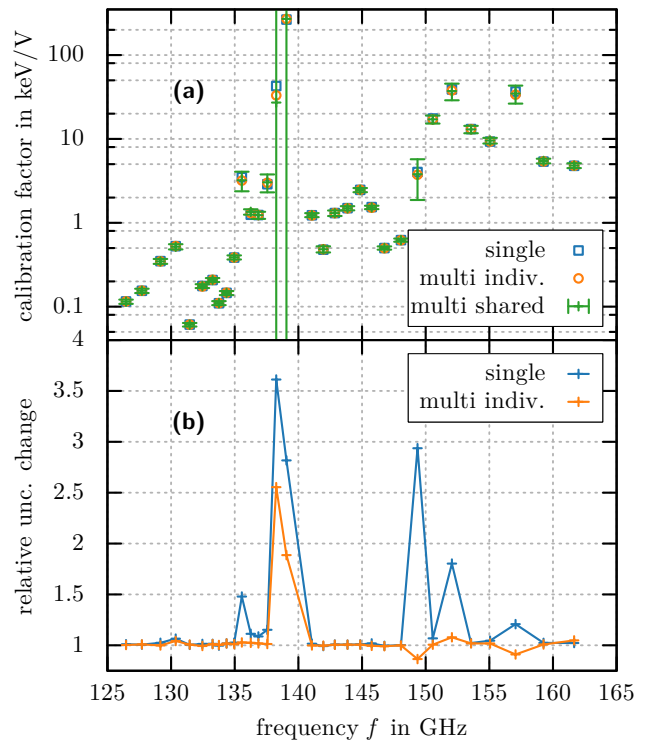


FIG. 11. (a) shows the frequency dependent calibration factors as obtained for single channel evaluations and the multi-channel analysis with and without a shared free parameter for the beam width. The uncertainties are only shown for the multi-channel analysis with a shared (and appropriately scaled) beam width, as this is the reference value for the relative uncertainty changes shown in (b). There, the standard deviation of the specified models is normalised by the standard deviation of the multi shared model. The given uncertainties correspond to one standard deviation as calculated from the MCMC samples. One can see that the single channel evaluation has the largest uncertainties, as expected. For channels with low sensitivities the reductions in the uncertainties that occur when switching from single channel evaluations to multi channel evaluations can be significant, as the more sensitive channels provide information about  $\phi_1^{\text{geo}}$  and  $\phi_2^{\text{geo}}$ . Using a common  $w$  reduces the uncertainties further, but less drastically.

expected to play an important role as well.

Figure 11b shows the uncertainties of the different models, normalised to the multi shared model. Going from the single channel evaluation to a combined model yields substantial decreases in the calibration factor uncertainties of insensitive channels, although the uncertainties for these channels remain very large. This phenomenon is most likely caused by the additional information about the geometrical properties  $\phi^{\text{geo}}$  that is mainly provided by stronger channels, helping the less sensitive channels to determine the begin and end of the hot/cold phases. Using a single beam parameter leads for a few channels to a small shift of the calibration factor, also reducing the uncertainties slightly, but less drastic than the switch from the single channel evaluation to a multi channel evaluation model.

Figure 12 shows the inferred intensity Gaussian beam width for each ECE channel. The beam width has been measured

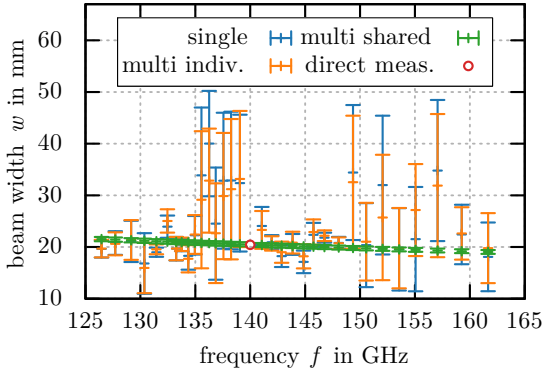


FIG. 12. The inferred Gaussian beam width of the microwave antenna characteristic for each channel. The direct measurement at 140GHz was done with a 440 point 2D scan of the beam intensities roughly 37.5cm away from the minimum beam waist.

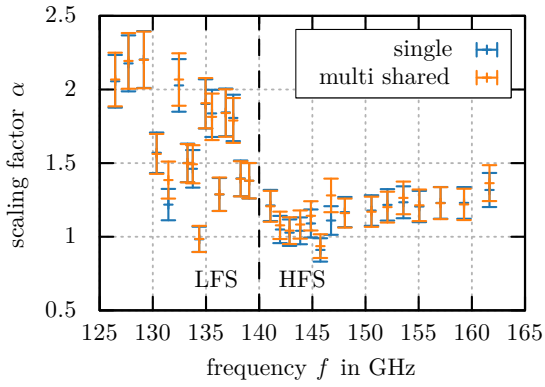


FIG. 13. The inferred variance scaling factors  $\alpha$  for the covariance is a measure of the model uncertainty for each channel. For the 'multi individual' case there is no visible difference to the 'multi shared' case, thus it has been omitted for the sake of clarity.

in the lab with a 140GHz source attached at the receiver end of the antenna to be roughly  $(20.4 \pm 0.2)$ mm at a distance of 37.5 cm away from the minimum beam waist, inferred with 10 million iterations of an MCMC of a simple Gaussian squared forward model that adapts the prediction uncertainties as well. This does not exactly match the distance at which the microwave foam is located relative to the minimum beam waist (which would be roughly 26cm to 32cm), but due to the small divergence of the beam width the introduced error is small. One can see that switching from the single channel analysis to a combined channel analysis slightly decreases the uncertainties for some less sensitive channels. In most cases the beam width shifts slightly towards values closer to the directly measured width. If only a single beam width, scaled according to equation 3, is used, the uncertainties get drastically reduced. The measured width is roughly 40 micrometer away from the predicted value by the model with a single beam parameter, with prediction uncertainties on the order of 0.4mm.

The variance scaling factors for the different channels are shown in figure 13. One can see that the values are not too far

away from 1, which indicates that the most relevant physic effects are considered. These values were reduced from values typically around 2.6 at the begin of the first W7-X experimental campaign by two changes of the setup: i) the horn was changed to include all polarizations and ii) the container for the liquid nitrogen had a round aperture that was changed to a rectangular aperture. The round aperture system was more sensitive to misalignments of the mirror-antenna system. Indeed an offset of around 3.5cm of the beam on the mirror was measured for the calibration used for the first experimental campaign, but the asymmetry introduced in the signal was not visible to the naked eye. As the physics for each channel should be similar,  $\alpha$  should have similar values for the different channels. A notable difference between the  $\alpha$  values is an indication that different physics effects play a role, or at least that these different effects are of different importance. One could expect that for channels with a small sensitivity electronics effects are of larger importance, therefore changing the variance scaling. Notable differences between the models that couple the beam width of different channels directly and those that do not, and between the individual and multi-channel evaluations are not observed. The values on the low field side (below 140GHz) scatter more and tend to be larger. A potential source for this behaviour can be found in hardware issues, respectively, implicit assumptions that are violated more strongly for low field side channels, although currently no such problem is known to the authors.

In summary, one can see that the single channel evaluation is satisfactory in most cases. However, if time is not a critical factor it is still beneficial to use a model combining the ECE channels.

## B. ECE Spectra

From the calibration procedure radiation temperature spectra can be derived from the measurements done during a plasma discharge. An example is shown in figure 14. The data originates from a 3.8s long plasma discharge that was heated on axis with electron cyclotron resonance heating<sup>34</sup> (ECRH). The ECRH power was 2.5MW in the first phase and was increased to roughly 5MW shortly after pellet fuelling started. The line averaged electron density as measured by a single channel dispersion interferometer<sup>35</sup> rose during pellet fuelling up to about  $7 \times 10^{19} \text{ m}^{-2}$ .

Channel 15 (138.26GHz) and 16 (139.06GHz) show a very low sensitivity, leading to radiation temperatures above 20keV and uncertainties of several hundred percent. Consequently, they were omitted in this plot. For channel 16 this is expected, as the channels frequency band locates it in the slope of the notch filter. Above 155GHz contributions from the third harmonic X mode emission start to play a role. The reconstruction of the electron temperature profile from the spectrum will be discussed in a separate publication. Figure 15 shows a comparison of the ECE timetrace of a channel close to the core to a Thomson scattering channel close to the core. The deviations in the first second are probably caused by the filters of the Thomson scattering system that lead to big

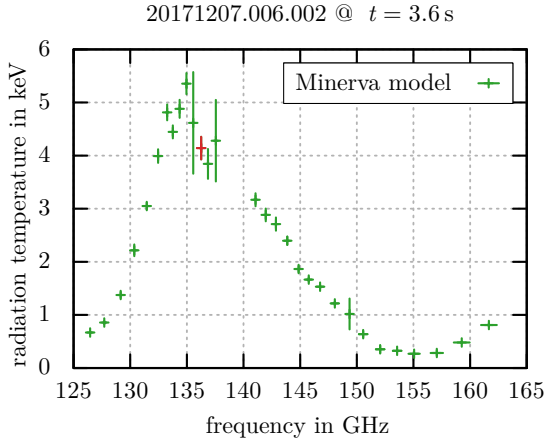


FIG. 14. Exemplary radiation temperature spectrum from an on axis ECRH W7-X plasma discharge, calculated with the calibration factors as obtained from the Minerva model that incorporates variance scaling. The channel marked with red corresponds to the timetrace shown in figure 15.

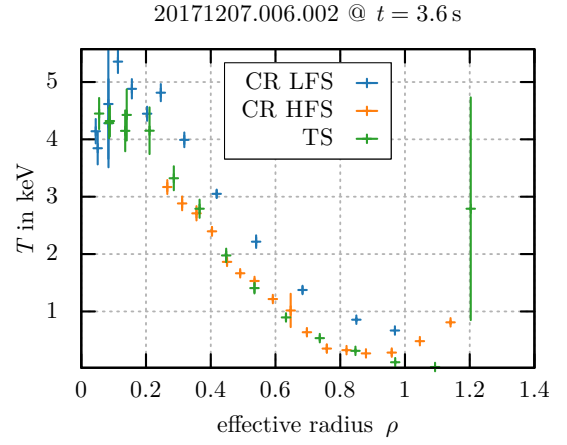


FIG. 16. Comparison of an ECE cold resonance profile with a profile from Thomson scattering. For a detailed discussion see section IV.

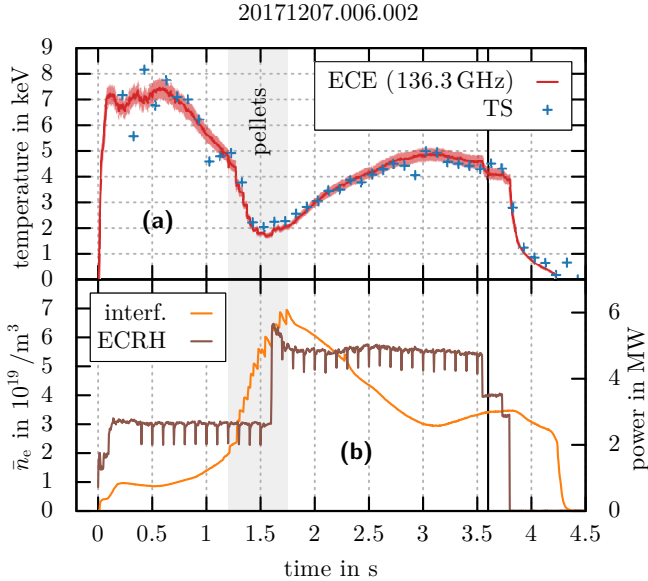


FIG. 15. **(a)** shows the calibrated signal of an ECE channel close to the core, compared with a central channel from the Thomson scattering system<sup>29</sup> for the plasma discharge also shown in figure 14. The line at 3.6s indicates the spectrum shown in figure 14. **(b)** depicts the ECRH and line averaged density from the single channel dispersion interferometry. The ECRH blips are necessary for the collective Thomson scattering diagnostic<sup>33</sup>.

uncertainties for electron temperatures above around 7 keV. The remaining seconds of the discharge show a good agreement for a wide range of electron densities and multiple power levels. Note that the Shafranov shift is small in W7-X<sup>36</sup>.

A comparison of the ECE radiation temperatures mapped to the cold resonance with Thomson scattering diagnostic  $T_e$  is shown in figure 16. One can see that the radiation temperatures on the high field (inboard) side and the electron temperatures measured by the Thomson scattering system agree

well, while the values on the low field side deviate notably. However, this can be caused by many factors; For example, the equilibrium chosen has a large influence on the positioning of the ECE, as the magnetic field gradient along the line of sight is small, such that small changes of the magnetic flux can change the position of the ECE channels drastically. Furthermore, there is no automatic equilibrium reconstruction at W7-X, so the equilibrium chosen here is determined automatically from a precalculated number of reference equilibria and not ideal for the selected time point. In addition, the emission of the ECE stems, if radiation transport is taken into account, from behind the cold resonance, which would move the low field (outboard) side ECE channels closer to  $\rho = 0$ . Finally, the Thomson scattering diagnostic does not include systematic uncertainties from their calibration procedure, so these uncertainties represent a lower threshold. Therefore, one cannot draw definite conclusions about the quality of the ECE calibration from the comparison to the  $T_e$  values measured by Thomson scattering.

## V. CONCLUSIONS

The Bayesian Minerva model developed for the multi channel consistent calibration of a generic microwave radiometer provides the sensitivities in an automated fashion, insensitive to signal drifts on timescales  $> 1$  s. While a classical analysis approach will yield comparable results (except for weak channels, for which the models in this paper will yield clearly better results), the approach presented in this paper lays out the underlying assumptions more clearly, makes it more clear to which extent the calibration procedure is understood and provides a measure of the beam width. Also, classical error propagation is linear while a Bayesian analysis can handle nonlinearities.

Moreover, it allows to obtain non-Gaussian posterior distributions for the calibration factors, although for the sensitive channels in the radiometer studied a Gaussian distribution provides an excellent fit to the posterior distribution.

The ECE spectra obtained from typical W7-X plasmas do not show unexpected or unphysical features. The uncertainties are for reasonably sensitive channels typically on the order of 6%. While the high dimensional (98D) model provides in principle the highest consistency, a comparison to the much quicker and parallelized single channel evaluation (4D for each channel) yields only little differences, justifying the use of the simpler approach.

Further improvements could be achieved by applying neural networks, as described in Pavone et. al.<sup>25</sup>. It also would be interesting to use the calibration model in combination with a plasma model containing predictions for Thomson scattering and ECE, guaranteeing consistent calibration factors across different diagnostic systems with the models presented in this paper providing the tested and prepared ECE calibration branch.

## ACKNOWLEDGMENTS

The authors wish to thank Karsten Ewert for his technical work support. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

## Appendix: Temperature dependence of the prediction

For a specific channel  $i$  the predicted signal in bit is described via equation (9). Given a linear sensitivity, this equation can be rewritten to

$$f_i^{\text{pred}}(\varphi) = T_{\text{eff},i}(\varphi) \tilde{\eta}_i \quad (\text{A.1})$$

$$= \frac{\Delta s_i T_{\text{eff},i}(\varphi)}{T_{\text{eff}}^{\text{hot}} - T_{\text{eff}}^{\text{cold}}}. \quad (\text{A.2})$$

As we do subtract the offset of the measured signal, we end up with

$$f_i^{\text{pred}}(\varphi) - \bar{f}_i^{\text{pred}} = \Delta s_i \left( \frac{T_{\text{eff},i}(\varphi)}{T_{\text{eff}}^{\text{hot}} - T_{\text{eff}}^{\text{cold}}} - \underbrace{\left( \beta - \frac{\gamma}{n} \right)}_{\text{offset}} \right), \quad (\text{A.3})$$

where

$$\beta = \frac{T_{\text{eff}}^{\text{hot}}}{T_{\text{eff}}^{\text{hot}} - T_{\text{eff}}^{\text{cold}}},$$

$\gamma$ : number of entries for the cold vector,

$n$ : number of entries for the cold and the hot vector.

As all effective temperatures are between  $T_{\text{eff}}^{\text{cold}}$  and  $T_{\text{eff}}^{\text{hot}}$ , they can be expressed by  $\beta - \delta$  with  $\delta \in [0, 1]$ . Thus, the predicted

signal can be written as

$$f_i^{\text{pred}}(\varphi) - \bar{f}_i^{\text{pred}} = \Delta s_i \left( \beta - \delta - \left( \beta - \frac{\gamma}{n} \right) \right) \quad (\text{A.4})$$

$$= \Delta s_i \left( \frac{\gamma}{n} - \delta \right). \quad (\text{A.5})$$

Thus, the predicted signal does not depend on the absolute values of the effective temperature for a linear sensitivity.

## Appendix: Details to calibration model

This section will go into details of the nodes shown in figure 6. The chopper channel needs to be set to the ECE channel that records the chopper signal. The calibration segments needs to be set to the nanosecond that the calibration starts and the nanosecond at which it ended. The radiometer settings are stored on the database and are fetched from there for the channel specific datasource. The datasource does the fetching of metadata and necessary preprocessing like the conditional averaging described in section III B (which got a separate node in figure 6 for the sake of clarity).

$T_{\text{H}_2\text{O}}$  represents a normal distribution with a mean of 280 K and a standard deviation of 10 K. The mean value has been guessed, as the vapour will not have a temperature substantially above 0°C. To take into account that this guess is very uncertain, a conservatively estimated standard deviation of 10 K was assumed.

$\epsilon_{\text{H}_2\text{O}}$  represents a uniform distribution ranging from 0.01 to 0.03 with an initial guess of 0.02. Also, these values have been guessed.

$\epsilon_{\text{mirror}}$  represents a uniform distribution ranging from 0.01 to 0.03 with an initial guess of 0.02. Also, these values have been guessed.

$T_{\text{LN}_2}$  represents a normal distribution with a mean of 77.2 K, which is the temperature of boiling nitrogen, and a standard deviation of 0.5 K as the boiling temperature of liquid nitrogen is well known.

$T_{\text{eff}}^{\text{hot}}$  represents a normal distribution with a mean of 294.45 K, which is the temperature in the torus hall, and a standard deviation of 3.5 K as a conservative estimate of the uncertainties.

$T_{\text{eff}}^{\text{cold}}$  Corresponds to equation 5.

$\Delta T_{\text{eff}}$  is calculated by  $T_{\text{eff}}^{\text{hot}} - T_{\text{eff}}^{\text{cold}}$ .

$\varphi_1^{\text{eff}}$  is a uniform distribution ranging from 0 to  $2\pi/6$  with a starting value of 0.3832, which has been found empirically to be a good starting point.

$\varphi_2^{\text{eff}}$  is a uniform distribution ranging from  $10\pi/6$  to  $2\pi$  with a starting value of 5.95973, which has been found empirically to be a good starting point.

# bins is the number of bins used for the conditional averaging. If it is set to 0, which is the default, a reasonable number is calculated from the average number of data points recorded during one mirror rotation.

$\Delta s_i$  represents the bit dip and is a uniform distribution ranging from 0 to 50000 with a starting value determined dynamically for each channel from the datasource.

$w_i$  represents the beam width and is a uniform distribution ranging from 0.05 to 0.3 with a starting value of 0.15384615384. The values are in radians. The initial guess corresponds to about 20mm.

$\alpha_i$  represents the variance scaling and is a uniform distribution ranging from 0.01 to 1000 with a starting value of 2.

$b_i$  is the bit to volt factor and is a normal distribution with a mean of  $3.71333 \times 10^{-4}$  and a standard deviation of  $3.71333 \times 10^{-6}$ . This values stem form separate measurements.

$\lambda_i$  is the factor that corrects the systematic differences between the invessel and the calibration optical system, see figure 2. It is a normal distribution with the mean at the channel specific measured correction factor and corresponding standard deviation.

$1/\tilde{\eta}_i$  corresponds to equation 6.

$1/\eta_i$  corresponds to equation 10.

$T_{\text{eff},i}(\varphi)$  corresponds to equation 7.

$f_i^{\text{pred}}(\varphi)$  corresponds to equation 9.

$f_i^{\text{meas}}(\varphi)$  corresponds to measured and conditionally averaged data (see the description of the datasource).

$\alpha_i \sigma_i^2$  corresponds to variance as estimated from the conditional averaging, multiplied by the variance scaling factor  $\alpha$ .

Lastly, the observation corresponds to a multivariate normal distribution with the mean set to the predicted values, a variance as given by  $\alpha_i \sigma_i^2$  and the actual value set to the observed values,  $f_i^{\text{meas}}(\varphi)$ .

- <sup>1</sup>R. H. Dicke, R. Beringer, R. L. Kyhl, and A. B. Vane, *Physical Review* **70**, 340 (1946).
- <sup>2</sup>W. A. Imbriale, S. Weinreb, and H. Mani, in *2007 IEEE Aerospace Conference* (2007).
- <sup>3</sup>E. Mazzucato, *Electromagnetic Waves for Thermonuclear Fusion Research* (World Scientific Publishing Company, 2014) Chap. 9.
- <sup>4</sup>H. J. Hartfuß and T. Geist, *Fusion Plasma Diagnostics with mm-Waves: An Introduction* (Wiley-VCH, Weinheim, 2013).
- <sup>5</sup>X. Liu, H. L. Zhao, Y. Liu, E. Z. Li, X. Han, C. W. Domier, N. C. L. Jr., A. Ti, L. Q. Hu, and X. D. Zhang, *Review of Scientific Instruments* **85**, 093508 (2014).
- <sup>6</sup>C. Beidler, G. Grieger, F. Herrnegger, E. Harmeyer, J. Kisslinger, W. Lotz, H. Maassberg, P. Merkel, J. Nührenberg, F. Rau, J. Sapper, F. Sardei, R. Scardovelli, A. Schlüter, and H. Wobig, *Fusion Technology* **17**, 148 (1990), <https://doi.org/10.13182/FST90-A29178>.
- <sup>7</sup>R. C. Wolf *et al.*, *Nuclear Fusion* **57**, 102020 (2017).
- <sup>8</sup>J. Svensson, A. Werner, and JET EFDA contributors, in *2007 IEEE International Symposium on Intelligent Signal Processing* (2007).
- <sup>9</sup>G. Taylor, M. E. Austin, A. Basile, J. H. Beno, S. Danani, R. Feder, S. Houshmandyar, A. E. Hubbard, D. W. Johnson, A. Khodak, R. Kumar, S. Kumar, A. Ouroua, S. B. Padasalagi, H. K. B. Pandya, P. E. Phillips, W. L. Rowan, J. Stillerman, S. Thomas, V. S. Udintsev, G. Vayakis, M. Walsh, and D. Weeks, *EPJ Web of Conferences* **147**, 02003 (2017).
- <sup>10</sup>R. Fischer, C. J. Fuchs, B. Kurzan, W. Suttrop, and E. W. and, *Fusion Science and Technology* **58**, 675 (2010).
- <sup>11</sup>S. K. Rathgeber, L. Barrera, T. Eich, R. Fischer, B. Nold, W. Suttrop, M. Willensdorfer, E. Wolfrum, and the ASDEX Upgrade Team, *Plasma Physics and Controlled Fusion* **55**, 025004 (2013).
- <sup>12</sup>S. S. Denk, R. Fischer, O. Maj, E. Poli, J. K. Stober, U. Stroth, B. Vanovac, W. Suttrop, M. Willensdorfer, and the ASDEX Upgrade Team, *European Physical Journal Web of Conferences* **147**, 02002 (2017).
- <sup>13</sup>S. S. Denk, R. Fischer, H. M. Smith, P. Helander, O. Maj, E. Poli, J. Stober, U. Stroth, W. Suttrop, E. Westerhof, M. Willensdorfer, and the ASDEX Upgrade Team, *Plasma Physics and Controlled Fusion* **60**, 105010 (2018).
- <sup>14</sup>S. Schmuck, J. Svensson, L. Figini, T. Jonsson, J. Fessey, L. Meneses, J. Boom, and JET EFDA Contributors, “Electron temperature and density

inferred from JET ECE diagnostics,” (2014), european Physical Society, 41<sup>st</sup> Conference on Plasma Physics.

- <sup>15</sup>S. Schmuck, J. Svensson, E. de la Luna, L. Figini, T. Johnson, B. Alper, M. Beurskens, J. Fessey, T. Gerbaud, A. Sirinelli, and JET EFDA Contributors, “Bayesian derivation of electron temperature profile using JET ECE diagnostics,” (2011), european Physical Society, 38<sup>th</sup> Conference on Plasma Physics.
- <sup>16</sup>M. Hirsch, U. Höfel, J. W. Oosterbeek, N. Chaudhary, J. Geiger, H.-J. Hartfuss, W. Kasperek, N. Marushchenko, B. van Milligen, B. Plaum, T. Stange, J. Svensson, H. Tsuchiya, D. Wagner, G. McWeir, R. Wolf, and the W7-X Team, *European Physical Journal Web of Conferences* (2019).
- <sup>17</sup>S. Schmuck, H.-J. Hartfuss, M. Hirsch, and T. Stange, *Fusion Engineering and Design* **84**, 1739 (2009), proceeding of the 25th Symposium on Fusion Technology (SOFT-25).
- <sup>18</sup>D. Wagner, W. Kasperek, F. Leuterer, A. Marcor, F. Monaco, M. Munich, E. D. Rijk, H. Schutz, J. Stober, M. Thumm, and A. Zeitler, in *2015 40th International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz)* (IEEE, 2015).
- <sup>19</sup>C. Fuchs and H. J. Hartfuss, *Review of Scientific Instruments* **72**, 383 (2001), <https://doi.org/10.1063/1.1309005>.
- <sup>20</sup>J. L. Séguin, D. Molina, G. Giruzzi, M. Goniche, G. Huysmans, P. Maget, and M. Ottaviani, *Review of Scientific Instruments* **76**, 123501 (2005).
- <sup>21</sup>D. Sivia and J. Skilling, *Data Analysis*, 2nd ed. (Oxford University Press, 2006).
- <sup>22</sup>C. M. Bishop, “Graphical models,” in *Pattern Recognition and Machine Learning*, Information Science and Statistics (Springer-Verlag New York, 2006) Chap. 8, pp. 359–362, 1st ed.
- <sup>23</sup>J. Svensson, O. Ford, D. C. McDonald, A. Meakins, A. Werner, M. Brix, A. Boboc, M. Beurskens, and JET EFDA Contributors, *Contributions to Plasma Physics* **51**, 152 (2011).
- <sup>24</sup>M. J. Hole, G. von Nessi, D. Pretty, J. Howard, B. Blackwell, J. Svensson, and L. C. Appel, *Review of Scientific Instruments* **81**, 10E127 (2010).
- <sup>25</sup>A. Pavone, J. Svensson, A. Langenberg, N. Pablant, U. Hoefel, S. Kwak, R. C. Wolf, and the W7-X team, *Review of Scientific Instruments* **89**, 10K102 (2018).
- <sup>26</sup>R. Hooke and T. A. Jeeves, *Journal of the ACM* **8**, 212 (1961).
- <sup>27</sup>H. Haario, E. Saksman, and J. Tamminen, *Bernoulli* **7**, 223 (2001).
- <sup>28</sup>C. J. Geyer, in *Handbook of Markov Chain Monte Carlo*, Handbooks of Modern Statistical Methods, edited by S. Brooks, A. Gelman, G. Jones, and X.-L. Meng (Chapman and Hall/CRC, 2011) Chap. 1.
- <sup>29</sup>S. A. Bozhenkov, M. Beurskens, A. Dal Molin, G. Fuchert, E. Pasch, M. R. Stoneking, M. Hirsch, U. Höfel, J. Knauer, J. Svensson, H. Trimino Mora, and R. C. Wolf, *Journal of Instrumentation* **12**, P10004 (2017).
- <sup>30</sup>E. Parzen, *The Annals of Mathematical Statistics* **33**, 1065 (1962).
- <sup>31</sup>B. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability (Taylor & Francis, 1986).
- <sup>32</sup>X. Han, X. Liu, Y. Liu, C. W. Domier, N. C. Luhmann, E. Z. Li, L. Q. Hu, and X. Gao, *Review of Scientific Instruments* **85**, 073506 (2014).
- <sup>33</sup>J. van den Berg, I. Abramovic, N. J. Lopes Cardozo, and D. Moseev, *Review of Scientific Instruments* **89**, 083507 (2018), <https://doi.org/10.1063/1.5035416>.
- <sup>34</sup>Stange, Torsten, Laqua, Heinrich Peter, Beurskens, Marc, Bosch, Hans-Stephan, Bozhenkov, Sergey, Brakel, Rudolf, Braune, Harald, Brunner, Kai Jakob, Cappa, Alvaro, Dinklage, Andreas, Erckmann, Volker, Fuchert, Golo, Gantenbein, Gerd, Gellert, Florian, Grulke, Olaf, Hartmann, Dirk, Hirsch, Matthias, Höfel, Udo, Kasperek, Walter, Knauer, Jens, Langenberg, Andreas, Marsen, Stefan, Marushchenko, Nikolai, Moseev, Dmitry, Pablant, Novomir, Pasch, Ekkehard, Rahbarnia, Kian, Mora, Humberto Trimino, Tsujimura, Toru, Turkin, Yuriy, Wauters, Tom, and Wolf, Robert, *EPJ Web Conf.* **157**, 02008 (2017).
- <sup>35</sup>K. Brunner, T. Akiyama, M. Hirsch, J. Knauer, P. Kornejew, B. Kursinski, H. Laqua, J. Meineke, H. T. Mora, and R. C. Wolf, *Journal of Instrumentation* **13**, P09002 (2018).
- <sup>36</sup>R. C. Wolf, S. Bozhenkov, A. Dinklage, G. Fuchert, Y. O. Kazakov, H. P. Laqua, S. Marsen, N. B. Marushchenko, T. Stange, M. Zanini, I. Abramovic, A. Alonso, J. Baldzuhn, M. Beurskens, C. D. Beidler, H. Braune, K.-J. Brunner, N. Chaudhary, H. Damm, P. Drewelow, G. Gantenbein, Y. Gao, J. Geiger, M. Hirsch, U. Höfel, M. Jakubowski, J. Jelonek, T. Jensen, W. Kasperek, J. Knauer, S. B. Korsholm, A. Langen-

berg, C. Lechte, F. Leipold, H. Trimino Mora, U. Neuner, S. K. Nielsen, D. Moseev, H. Oosterbeek, N. Pablant, E. Pasch, B. Plaum, T. Sunn Peder-

sen, A. Puig Sitjes, K. Rabbarnia, J. Rasmussen, M. Salewski, J. Schilling, E. Scott, M. Stejner, H. Thomsen, M. Thumm, Y. Turkin, F. Wilde, and the W7-X team, *Plasma Physics and Controlled Fusion* **61**, 014037 (2018).