Parametrization of reflectometry fluctuation frequency spectra for systematic study of fusion plasma turbulence

Y. Sun,[1,2,3,a)] R. Sabot,[1] G. Hornung,[3] S. Heuraux,[2] S. Hacquin,[1,4] and G. Verdoolaege[3,5]

[1)] *CEA, IRFM, F-13108 Saint Paul Lez Durance, France*

[2)] *IJL UMR 7198 CNRS, Université de Lorraine, F-54000 Nancy, France*

[3)] *Department of Applied Physics, Ghent University, 9000 Gent, Belgium*

[4)] *EUROfusion Programme Management Unit, Culham Science Centre, Culham, OX14 3DB, UK*

[5)] *LPP-ERM/KMS, B-1000 Brussels, Belgium*

We describe a way to parameterize power spectra extracted from fixed-frequency reflectometry data, with a view to systematic studies of turbulence properties in tokamak plasmas. Analysis of typical frequency spectra obtained from a new database suggests a decomposition in a set of four key components: the direct current (DC) component, low-frequency (LF) fluctuations, broadband (BB) turbulence and the noise (N) level. For the decomposition in the identified components, different kinds of functions are tested and their fitting performance is analysed to determine the optimal spectrum parametrization. In particular, for the BB turbulence three models are compared qualitatively based on a number of representative spectrum test cases, notably the generalized Gaussian, the Voigt and the Taylor model. In addition, quantitative performance testing is accomplished using the weighted residual sum of squares (RSS) and the Bayesian information criterion (BIC) in a large database including $350,000$ spectra obtained in Tore Supra. Next, parametrization by the Taylor model is applied to Ohmically heated plasmas, and a BB energy basin is systematically observed in the core plasma region, which shrinks with decreasing radial position of the $q = 1$ surface. This basin might be explained by a drop of the density fluctuation level inside the $q = 1$ surface.

[a)] Electronic mail: yan.sun@cea.fr

I. INTRODUCTION

Turbulence is the cause of anomalous transport, degrading plasma confinement in tokamaks, of which turbulent density fluctuations form an important aspect. Among many turbulence diagnostics[1], reflectometry[2] is an extensively used method due to its convenient implementation and high spatial resolution. The properties of density fluctuations have been studied by reflectometry experimentally[3] in tokamaks for decades and, following improvements of interpretative models deduced from reflectometry simulations[4], significant knowledge of turbulence properties in tokamak plasmas has been obtained.[5] Specifically, great efforts have been made to measure the turbulence level, owing to its direct link to the transport coefficients. It was found that the turbulence level reaches a maximum in the plasma edge, becoming lower in the core region.[6,7] The power spectrum from density fluctuations contains abundant information about different kinds of plasma instabilities, making it a powerful tool to study transport and confinement in fusion plasmas. For instance, zonal flows (ZF) and geodesic acoustic modes (GAM) can be observed in power spectra, reflecting the complicated impact of turbulence on the transport properties.[8] Furthermore, MHD modes, including many kinds of Alfvén modes[9], are important for steady-state operation as well as advanced tokamak configurations. Moreover, low-frequency and high-frequency quasi-coherent (QC) oscillations have been identified and linked to different drift wave instabilities.[10–13]

In the past, however, experimental analyses have concentrated on limited numbers of discharges with selected parameters. In this traditional approach, only one or a few parameters are allowed to change, while others are kept constant. This way, analyzing trends occurring over a large range of the operational space, with multiple variables changing at the same time, is difficult. As a complementary approach to more common studies based on a limited, well-controlled data set, we propose a systematic analysis using a large database obtained under a wide variety of plasma conditions. The possibility to detect important structure in the database (trends and clusters) depends crucially on an efficient *parametrization* of the turbulence properties extracted from the data. The parametrization method relies on a decomposition of the turbulence spectrum. This decomposition method is based on the pioneering works of Vershkov *et al.*[14] and Krämer-Flecken *et al.*[15] Through radial, poloidal and long-range correlation, they have investigated different components of the turbulence

spectrum and their properties. Our work is based on such an approach, using an extensive database of Tore Supra plasmas and a parametrization of frequency spectra obtained from reflectometry measurements. This paper focuses on the methodology leading towards that parametrization.

In order to link the properties of the fluctuation spectrum to the underlying physics in a large set of discharges under widely varying plasma conditions, it is very important to extract in a robust way the salient features from the data and describe them using a minimal set of parameters. In this work, three different parametric models were used and their performance over the entire database was compared quantitatively. Great care was taken to obtain the optimal solution of the parametrization, taking into account the physical interpretation of the various spectrum components, with satisfactory results. The parametrization method is then applied to investigate the evolution of turbulence w.r.t the edge safety factor (q_ψ), one of the dimensionless parameters. A reduction of the broadband component of turbulence spectrum was observed in the core and the localization of the reduction seems to be linked to the sawteeth instabilities.

The rest of this paper is organized as follows. In Sec. II, first the fixed-frequency reflectometry diagnostic setup is briefly introduced. Next, we discuss decomposition and parametrization of the frequency spectra. Section III is dedicated to the parametrization results and comparison of the three models. The application to Ohmically heated plasmas is shown and discussed in Sec. IV. Finally, conclusions and perspectives of future work are given in Sec. V.

II. SYSTEMATIC STUDY OF TURBULENCE PROPERTIES

A. Core reflectometry diagnostic

A D-band heterodyne reflectometry diagnostic was operated on Tore Supra from 2002 to the last discharge in 2011. It covered the frequency range from 100 to 155 GHz in X-mode and was designed to probe the core from mid-radius on the low-field side (LFS) to the high-field side (HFS) at high magnetic field (3.5 $\sim$ 3.8 T). The lower-field (3 $\sim$ 3.5 T) measurements were strongly restricted in their radial positions, with limited accessibility to the LFS. The reflectometer was located on the equatorial plane. Low-divergence (HPHW

$\sim 1^{\circ}$) antennas (bistatic setup) produced an almost parallel beam spot about 3.5 cm in radius.[16] Heterodyne detection was provided by a single sideband modulator setup.[17] A second independent hopping channel (fixed frequency steps) was added in 2005 to probe two radial positions simultaneously and to perform radial correlation measurements. The reflectometer usually probed 2×20 fixed frequency steps of 10 milliseconds several times per shot. The overall measurement usually lasted less than 300 ms during stable phases of the discharge. Long acquisitions (500 ms) with > 100 ms plateaus could also be performed for MHD studies.[18] The hopping acquisition sampling frequency was 1 MHz, with typically 10,000 data points in 10 ms. A heterodyne I/Q detection allowed to get the complex signal:

$$s(t) = A \cdot (cos\phi + i \cdot sin\phi),$$

where $A = A(t)$ and $\phi = \phi(t)$.

In carrying out the fast Fourier transform (FFT), the number of frequency bins was set to 1024 in order to obtain good resolution at low frequency. The Welch's method[19] was used to calculate the spectra with Hamming window and 50% overlap to reduce sidelobes. Each frequency spectrum was thus given by the typically averaging over about 20 spectra.

One of the two reflectometry channels is also equipped with a voltage controlled source[16] much faster than the frequency synthesiser employed for fluctuation measurements. This source was used for measuring typically 1000 density profiles per discharge. Each profile was acquired within brief time windows of 40 to 100 μs. As the profile and frequency sources are coupled to the same millimeter hardware, profiles and fluctuations cannot be measured at the same time.

B. Decomposition of fluctuation frequency spectra

Figure 1 shows some typical frequency spectra obtained from fluctuation measurements using this reflectometry setup under different conditions and at varying radial positions. The spectra $S(f)$ (f is frequency) are plotted on a logarithmic scale ($10 \times log_{10}(S)$ in units of decibel (dB)). Although not all possible shapes of the complicated and varying frequency spectra in Tore Supra plasmas are shown, the examples in Fig. 1 do represent the typical spectral shape features encountered throughout the database. Positive and negative frequencies correspond to positive and negative directions, respectively. The spectrum can be

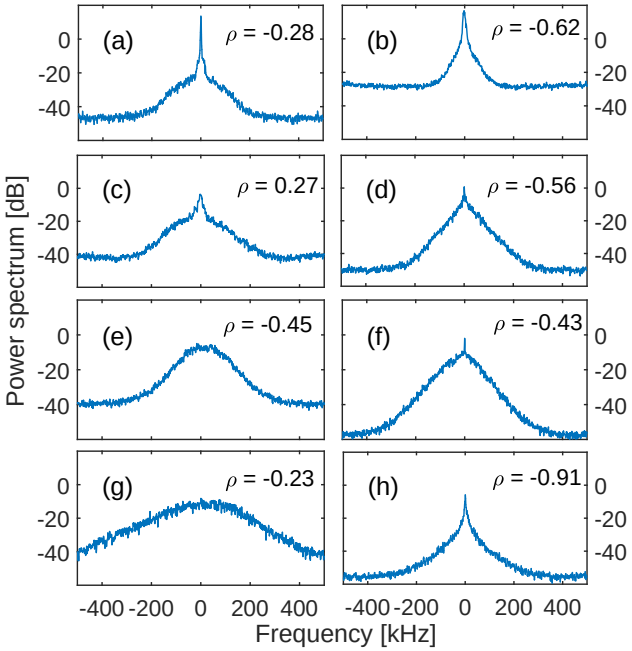

Figure 1. Some typical frequency spectra obtained from the core reflectometer database, with 1024 frequency bins. The Welch method was used with Hamming window width 1024 and an overlap of 50%. The normalized radial position (ρ) of the cutoff layer was calculated from a density profile obtained by the interferometry diagnostic. Negative ρ indicates HFS radial position.

almost symmetrical, but sometimes the asymmetry is strong. This can be due to various reasons, like the Doppler shift, small displacements of the plasma with respect to the equatorial plane, or asymmetries of the turbulent structures or in the wave propagation. As shown in Fig. 1, the fluctuation frequency spectra can be Gaussian-like (spectra (e) and (g)) or much more Lorentzian-like (spectrum (h)), i.e. strongly peaked with heavy tails. Other spectrum shapes are in between these typical spectra. The low-frequency component can be intense (spectra (a) and (b)), invisible (spectra (e) and (g)), or mixed with other parts of the spectrum (spectra (d) and (h)).

Although manually investigating individual spectra by quantifying their properties (e.g. energy and width) on a case-by-case basis is possible, systematic and standardized investigation of numerous spectra with complicated shapes requires automated methods. Due to the often complex spectrum shape and in order to link various aspects of the shape to the underlying physical mechanisms, we decided on an approach wherein the spectrum is decomposed in several components. Every component is characterized by only a few param-

eters, hence the important features of all spectra in the database are described by a limited number of parameters, facilitating systematic studies.

In previous works[11,12,14,15,20] several components were distinguished in fluctuation spectra associated to specific physical phenomena: the direct current (DC) component[15], low-frequency (LF) fluctuations[14], broadband (BB) fluctuations and in some cases quasi-coherent (QC) oscillations[20]. The BB fluctuations, which cover the whole frequency range, have a short correlation length[14,15] and have been attributed to turbulence, to be called BB turbulence[12] hereafter. Both the LF and QC components are superimposed on the BB turbulence. The (LF) component represents the more intense fluctuations at low frequencies. Zonal flows and certain MHD modes like sawteeth could contribute to this component, but this is beyond the scope of this study. In addition, a very narrow central spike at zero frequency was identified as the reflectometer carrier wave, named the DC component in Ref. 15. The QC oscillations can be observed in the LFS and are linked to drift wave instabilities.[13,14] In addition, the noise (N) level should be considered as another component for completeness.

The central idea of our work is that, under the condition that the decomposition provides a faithful representation of the various spectrum components, and assuming that the main contribution to the density fluctuations originates from the vicinity of the cut-off layer, systematic studies of the underlying physical phenomena and their coupling should become feasible.

C. Parameter reduction and criteria

As mentioned before, to enable systematic studies of trends or evolution of turbulence properties, it is important to describe the spectrum components using a limited number of parameters. This is accomplished by modeling each component by a simple parameterized function, which is able to represent the shape of the component under different physical conditions. Thus, the objective is to fit the frequency spectrum by a model $S_{fit}(f)$, written as a sum of m components $C_i(f)$ $(i = 1, \ldots, m)$.

While the total number K of parameters describing the spectra should be limited for systematic studies and also to avoid overfitting, we still wish to cover all spectrum shapes observed in the database. Hence a moderate K should be aimed for. Further criteria for evaluating the fit quality are *flexibility*, *discrimination* and *robustness*. Flexibility refers

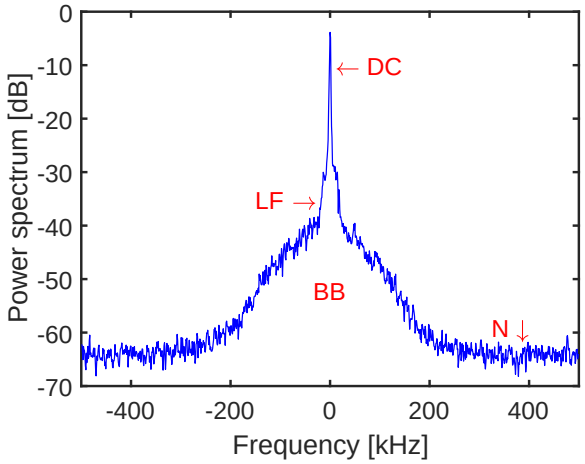

Figure 2. Typical spectrum (Fig. 1 (a)) with the various components to be fitted. The spectrum has been normalized to its total power.

to the ability of the model to represent many different spectral shapes, as seen in Fig 1. Discrimination is related to the distinguishing power of the model parameters w.r.t. the different spectral shapes, in the sense that the parameters should have moderate sensitivity to the spectral shape. Concretely, low sensitivity generates the same parameter results for all the cases, while high sensitivity leads to unstable parameters. Finally, robustness means that the parameters should have minimal dependence on small model deviations that are of little interest to the analysis, like noise.

D. Fitting methodology

From Fig. 1 it is clear that the power of the reflected signal can vary significantly (several dB or even more) from one spectrum to another, which can be attributed to multiple reasons. Specifically, the launched microwave power changes with wave frequency and the reflected microwave power decreases with deeper penetration. However, the absolute value of each parametric spectrum component should be comparable across spectra to allow systematic investigations. Therefore, *normalization* of the power spectrum is required and here the spectrum $S(f)$ is normalized to the integrated power:

$$\hat{S}(f) = \frac{S(f)}{\int_{f_{min}}^{f_{max}} S(f)\,\mathrm{d}f}. \tag{1}$$

Here, f_{min} and f_{max} denote the minimum and maximum frequency in the spectrum, which here we fix at $f_{min} = -500$ kHz and $f_{max} = 500$ kHz. As a result, the normalized spectrum integrates to unity, allowing spectra to be compared conveniently. For simplicity, in the remainder the normalized spectrum is also denoted by $S(f)$. All spectra used in this study were normalized in this way before parametrization in Sec. III as well as the turbulence analysis thereafter.

Figure 2 shows a typical normalized spectrum (corresponding to Fig. 1 (a)), with the various components indicated. When fitting a spectrum by minimization of the residual sum of squares (RSS), it is important to consider the scale at which to perform the fit. If the logarithmic scale is used, the results will be more representative for the highest frequencies, while on a linear scale the fit will tend to match primarily the low-frequency parts. For this reason a combination of fitting on both the logarithmic and linear scales is performed, by minimizing the following cost function:

$$F_{cost} = (1 - w) \times \frac{|\lg(S_{fit}) - \lg(S)|^2}{A_{lg}} + w \times |S_{fit} - S|^2. \tag{2}$$

Here, $S_{fit} = S_{fit}(f)$ and $S = S(f)$ denote the fitting model and the normalized frequency spectrum, respectively. In addition, $A_{lg} = \int_{Fmin}^{Fmax} (lg(S))^2 \, \mathrm{d}f$, where $lg = 10 \times log10$, is the integral of the spectrum on the logarithmic scale, ensuring normalization of the logarithmic part of the cost function. As a result, it is possible to weigh the two parts of the cost function using a weight factor w, allowing a more proper fit of both the high-frequency and low-frequency parts of $S(f)$. We have chosen to give equal weight to the linear and logarithmic parts ($w = 0.5$). Experimentation with other values (0.25 and 0.75) has pointed out that the results are not very sensitive to the weight factor. This does not exclude a more optimal weight factor for different spectrum decompositions or different databases.

Apart from the components mentioned above, various low-frequency MHD modes (e.g. sawteeth, fishbones, tearing modes) and other high-frequency fluctuations (e.g. GAMs, beta Alfvén eigenmodes, toroidal Alfvén eigenmodes) could appear under certain conditions. Since the bandwidth of these fluctuations is relatively narrow, the contribution to the total power can safely be neglected, even though their amplitudes can be large in some cases. The fitting results are therefore not expected to be substantially influenced in the presence of such modes. On the other hand, QC oscillations can attain significant bandwidths (tens of kHz). Low-frequency and high-frequency QC modes have been observed and examples can

be found in Refs. 11–14, and 20. However, we did not consider QC modes in the present stage, as their contribution to the power on the logarithmic scale is limited anyway.

In summary, every spectrum is decomposed into four basic components: the direct current (DC) component, the low-frequency (LF) fluctuations, the broadband (BB) turbulence and the noise (N) level, as shown in Fig. 2. Therefore the number of components m is 4:

$$S_{fit} = C_{DC} + C_{LF} + C_{BB} + C_N. \tag{3}$$

E. Fitting functions

The next step in the spectrum decomposition is to select the appropriate fitting functions for each component.

1. The noise level

The level of noise, assumed to be frequency-independent white noise, can be described by a single constant, therefore

$$C_N = \epsilon_N(f). \tag{4}$$

The noise level is also helpful to identify trivial spectra with low signal-to-noise ratio (SNR).

2. The low-frequency parts

The low-frequency parts of the spectrum include the DC and LF components. For each component, we need at least three parameters to describe the intensity, the central position and the spectral shape. Inspired by the normalization to unity of the total spectrum, we choose various probability density functions (PDFs) to model each of the components. The Gaussian (normal) PDF is the most straightforward choice, which has been used before as a model to describe the DC and LF components of coherence spectra.[15] In this case, the fitting functions for the DC and the LF components are

$$C_i = A_i \exp\left[-\frac{1}{2}\left(\frac{f-\mu_i}{\sigma_i}\right)^2\right], \tag{5}$$

where i is DC or LF. The amplitude A_i, the mean value μ_i and the standard deviation σ_i describe the intensity, central position and width of the components, respectively. For

more accurate fitting of the DC component, the zero frequency is placed at the center of the spectrum by using 1025 rather than 1024 frequency bins.

3. The BB turbulence

The Gaussian function was also considered for the BB turbulence, but was found insufficiently flexible to model all shapes. Indeed, the shape of the broadband can be distinctly non-Gaussian, more specifically Lorentzian (also known as Cauchy distribution) or Laplacian, with a strong peak and heavy tails, especially at the HFS. A combination of several Gaussian functions was tried as well, but that often caused the LF component to fit the BB instead. A more flexible function is therefore required and the following three options were explored: the *generalized Gaussian function* (GG), the *Voigt function* and the *Taylor function*, described below.

The BB turbulence using the generalized Gaussian function becomes

$$C_{BB}^{\mathbf{GG}} = A_{BB} \exp\left[-\left(\frac{|f-\mu_{BB}|}{\alpha_{BB}}\right)^{\beta_{BB}}\right], \tag{6}$$

where the fixed exponent in the Gaussian is replaced by a shape parameter β_{BB}, and the standard deviation $\sigma_{BB} = \sqrt{\alpha_{BB}^2\Gamma(3/\beta_{BB})/\Gamma(1/\beta_{BB})}$, describing the spectral width. This function can fit multiple shapes, like Gaussian ($\beta_{BB} = 2$) and Laplacian ($\beta_{BB} = 1$).

The Voigt function is a convolution of a Gaussian and a Lorentzian function:

$$C_{BB}^{\mathbf{Voigt}} = A_{BB} \int_{-\infty}^{+\infty} G(f;\sigma_{BBG}) L(\mu_{BB} - f;\gamma_{BBL})\,\mathrm{d}f, \tag{7}$$

where $G(x;\sigma)$ and $L(x;\gamma)$ are the centered (zero-mean) Gaussian and Lorentzian function, respectively, while μ_{BB} encodes the central position of the BB component. The Voigt function has been widely used for fitting spectral lines.[21]

A third alternative model for the BB component is the Taylor function. It was used in Ref. 22 to express the correlation function of a turbulence signal in plasma physics:

$$F_{corr}(k,u,\tau) = \exp\left[-k^2u^2\tau^2\left(\frac{t}{\tau} - 1 + e^{-t/\tau}\right)\right].$$

Here, k, u and τ represent the wavenumber, velocity and correlation time of the turbulence, respectively, while t is the sampling sequence, based on the theory of collective wave scattering by a non-uniform plasma.[23] Since the velocity correlation $C_v = e^{-t/\tau}$ was used, the

correlation function above is referred to as the Taylor function. The corresponding frequency spectrum is calculated through the Fourier transform of F_{corr}.

In scattering theory, long correlation lengths correspond to a convective Gaussian spectrum, while short correlation lengths have a diffusive Lorentzian spectrum. Accordingly, in F_{corr}, $ku\tau \gg 1$ and $ku\tau \leq 1$ lead to the convective and diffusive limit, respectively. Defining $\Delta = k^2 D = k^2 u^2 \tau$, where $D = u^2\tau$ is the diffusion coefficient, and introducing the averaged shift parameter $\delta\varphi = \exp(i2\pi\mu_{BB}t)$, the Taylor model of the BB turbulence has the following form:

$$\begin{aligned} C_{BB}^{\mathbf{Taylor}} &= A_{BB} \\ &\times \mathrm{FFT}\left\{\exp\left[-\Delta_{BB}(t - \tau_{BB} + e^{-t/\tau_{BB}})\right] \times \delta\varphi\right\}. \end{aligned} \tag{8}$$

The connection to the transport coefficients can then be established by analysis of the fitting parameters.

The number of parameters for the BB turbulence component is four, no matter which function is used. Together with the other three components, the complete fitting model S_{fit} has $K = 11$ parameters. Compared with the original 1024 frequency bins in the spectrum, the number of parameters has been reduced by two orders of magnitude. This opens the way to systematic investigations of the spectrum properties, which would have been very cumbersome with a large number of parameters.

F. Database

The database used for systematic studies in this work contains diagnostic characteristics, global operating parameters, local plasma parameters and the spectrum fitting parameters.

The diagnostic characteristics include reflectometry acquisition parameters, as well as the probing frequencies ($100 \sim 155$ GHz) and the radius of the cutoff layer. The radius of the cutoff layer is recovered from the density profile obtained from an interferometry diagnostic[24], as the density profile from the reflectometer is not available during fluctuation measurements. The normalized radius ρ of the cutoff layer ranges from -1 to 1, covering the entire plasma region. Here, $\rho = 0$ corresponds to the magnetic axis and negative values are used forthe HFS.

The global and local plasma parameters are obtained or calculated from various diagnostic data available in the Tore Supra database. Global operating parameters include the on-axis

toroidal magnetic field $B_{t,0}$, plasma current I_p, line-integrated electron density, major radius R, minor radius a, plasma heating power, elongation, edge safety factor (q_ψ) and more. Local plasma parameters include electron density n_e, electron temperature T_e, magnetic field B_t and scale length of the plasma permittivity (L_ϵ) at the cutoff positions.

The 11 fitting parameters from the fitting model S_{fit} reflect the turbulence properties. Among these, the two shape parameters of the BB turbulence are the most important.

So far only short time measurements (< 32 ms) are included in the study, under the assumption that the plasma parameters remain unchanged during the measurement period. Longer time measurements will be included in the future. Currently, the database contains more than 350,000 entries from about 6,000 discharges, mainly with Ohmic heating (OH), lower hybrid (LH) heating or ion cyclotron resonance heating (ICRH), and a limited number with electron cyclotron resonance heating (ECRH).

III. PARAMETRIZATION OF FREQUENCY SPECTRA

A. Constraints on component parameters

In order to maintain correspondence between each of the functional forms presented before and the spectrum components that they are intended to fit, additional constraints on the component parameters are necessary. To force the DC component to fit the narrow carrier wave at zero frequency, we impose the constraints $|\mu_{DC}| < 1$ kHz and $\sigma_{DC} < 2.5$ kHz, as 1 kHz is the frequency resolution of the spectrum. For the LF fluctuations, which sometimes include high- amplitude, low-frequency MHD modes up to a few kHz, the constraints are $|\mu_{LF}| < 10$ kHz and $\sigma_{LF} < 20$ kHz. Furthermore, to avoid overlap between the DC and LF components, we require $\sigma_{LF} > 1.5\,\sigma_{DC}$ and $\sigma_{LF} > 1$ kHz, where the factor 1.5 was determined empirically. To summarize, the constraints on the low-frequency part are:

$$\begin{aligned} &|\mu_{DC}| < 1 \text{ kHz}, \quad |\mu_{LF}| < 10 \text{ kHz}, \\ &\sigma_{DC} < 2.5 \text{ kHz}, \quad 1 \text{ kHz} < \sigma_{LF} < 20 \text{ kHz}, \\ &\sigma_{LF} > 1.5\,\sigma_{CS}. \end{aligned} \tag{9}$$

Constraints on the amplitudes and noise are not necessary.

For the BB turbulence, the constraints depend on the fitting functions. With the generalized Gaussian model, to separate the BB and LF components the constraints $\sigma_{BB} > 1.5\,\sigma_{LF}$

and $\sigma_{BB} > 10$ kHz are applied, where σ_{BB} is the standard deviation of the BB turbulence. In addition, to avoid an overly peaked BB fit, β_{BB} is assumed to be larger than 0.5, the generalized Gaussian function approximating a uniform distribution for large β_{BB} (in practice $\beta_{BB} > 8$).

For the Voigt model, no limits have been put on the Lorentzian part. As for the Gaussian part, we use the same constraints as in the generalized Gaussian model for the standard deviation σ_{BB}.

The parameters of the Taylor model are more difficult to constrain, as the two parameters Δ_{BB} and τ_{BB} jointly affect the spectral shape. Here, we set $\Delta > 0.01$ and $\tau_{BB} > 0.01$, to avoid unrealistically peaked shapes.

The constraints for the three models are summarized as follows:

- Generalized Gaussian model:
  $\sigma_{BB} > 10$ kHz, $\sigma_{BB} > 1.5\,\sigma_{LF}$, $0.5 < \beta_{BB} < 8$;
- Voigt model:
  $\sigma_{BBG} > 10$ kHz, $\sigma_{BBG} > 1.5\,\sigma_{LF}$;
- Taylor model: $\Delta_{BB} > 0.01$, $\tau_{BB} > 0.01$.

B. Optimization initial conditions

An interior-point algorithm was used for minimizing the cost function in (2). A more powerful global optimizer could be employed, but this turns out to be too time-consuming in practice for a database including 350,000 spectra. Therefore, multiple starting points were chosen based on various simple criteria, where the chance of converging to the global minimum increases with increasing number of initial guesses N_{iv}, striking a balance between computational load and goodness-of-fit.

For the DC component, A_{DC}, μ_{DC}, and σ_{DC} were estimated by the maximum value of the spectrum, and its first and second central moments in the frequency range $|f| < 3$ kHz, respectively. A similar approach was taken for the LF component, but within the frequency range 3 kHz $< |f| <$ 20 kHz to avoid influence by the strong DC component.

Likewise, for the BB component the parameters A_{BB} and μ_{BB} were estimated from the maximum and the first moment of the spectrum in the frequency range 20 kHz $< |f| <$

300 kHz, to avoid influence of the low-frequency components. The initialization of the other parameters depends on the model.

For the generalized Gaussian function, σ_{BB} and β_{BB} were estimated from the second central moment and standardized fourth moment (kurtosis), respectively. Multiple initial guesses were achieved by changing the starting β_{BB} (see below).

When fitting the BB turbulence by the Voigt function, calculation of the error function is time-consuming. The pseudo-Voigt function provides an approximation of the Voigt by using a linear combination rather than a convolution of the Gaussian and Lorentzian functions:

$$V_p(x) = \eta \cdot L(x, F) + (1 - \eta) \cdot G(x, F), \quad 0 \leq \eta \leq 1. \tag{10}$$

Here, F is the full width at half maximum (FWHM) and η is the weight coefficient between the two functions. The formula of F and η used is described in Ref. 25, where F and η are functions of σ_{BBG} and σ_{BBL}. The same constraints as for the Voigt function were used. The second moment of the spectrum gives the initial value of σ_{BBG} and multiple initial guesses of σ_{BBL} were obtained by varying η.

In the Taylor function, Δ_{BB} and τ_{BB} are slightly more difficult to estimate since they are not directly linked to the spectral shape. A tabulation of the standard deviation of $C_{BB}^{\mathbf{Taylor}}$ in (8) in terms of Δ_{BB} was made for $\tau_{BB} = 0.1$, allowing to derive initial estimates of Δ_{BB} from the second moment of the spectrum. Multiple initial guesses were realized by varying τ_{BB}. To determine the number of starting points N_{iv}, the generalized Gaussian model is taken as an example. The initial value of β_{BB} estimated from the kurtosis is denoted by β_{BB0} and was used as a first initial guess. Since $0.5 < \beta_{BB} < 8$ and in the database β_{BB} is typically between 1 and 2, the following initial values can cover the possible spectral shapes: $\beta_{BB0}/4$, $\beta_{BB0}/2$, $2\beta_{BB0}$, $4\beta_{BB0}$. For the second initial guess, the value of $\beta_{BB0}/4$ was used, because of all other initial values it corresponds to the shape differing the most from the shape associated with the first guess β_{BB0} of β_{BB}. This principle was also used to choose the third, fourth and fifth initial guess, i.e. $4\beta_{BB0}$, $\beta_{BB0}/2$ and $2\beta_{BB0}$, respectively. The convergence performance was evaluated through the averaged relative error of the overall fit for 1000 random spectra from the database, for different N_{iv}. From Fig. 3, the relative error is near 10% for a single initial value and decreases rapidly when increasing N_{iv} before saturation at $N_{iv} = 3$. At this point the relative error drops to $\sim 1.8\%$, meaning that the results are very close to the global minimum.

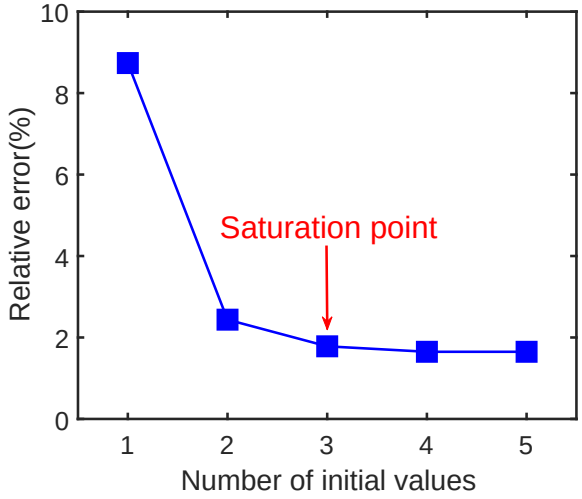

Figure 3. Relative error for the total fit, averaged over 1000 random spectra from the database, as a function of the number of initial values N_{iv} for β_{BB} in the GG model.

For the Voigt model, $\eta = 0$ and $\eta = 1$ denote the Gaussian and Lorentzian shape, respectively. Therefore the first and second initial guesses were obtained by setting $\eta = 0$ and $\eta = 1$, followed by three more initial values in between these extremes: $\eta = 0.5$, 0.25, 0.75. The relative fitting error saturates at around 7% beyond $N_{iv} = 5$. As for the Taylor model, empirical evaluation revealed a typical value of $\tau_{BB} = 0.1$. Therefore, we start from $\tau_{BB} = 0.1$ and then alternately increase and decrease according to the sequence $\tau_{BB} = 1$, 0.01, 0.5, 0.02. Again, the results remain almost the same for $N_{iv} > 5$, resulting in a fitting error of about 2%.

C. Comparison of models

1. Statistical comparison

To compare the fitting performance of the three models for the BB turbulence, the quality of the total fit was assessed for 10,000 spectra (about 3% of the full database). Spectra with low signal-to-noise ratio (SNR) were not considered for the analysis even though the fitting results are good. The performance was evaluated by means of the minimal value of the cost function ($F_{cost,min}$) and the Bayesian information criterion (BIC). Assuming a Gaussian

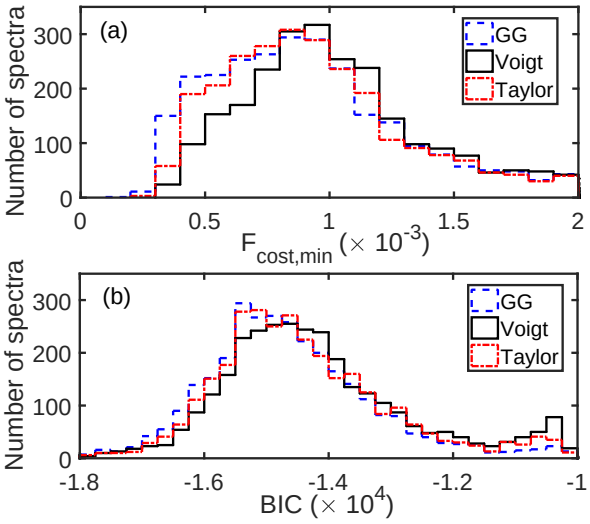

Figure 4. Comparison of (a) the minimal value of the cost function ($F_{cost,min}$) and (b) the Bayesian information criterion (BIC) for the generalized Gaussian (GG) function, the pseudo-Voigt function, and the Taylor function fitted to the BB turbulence component. The more a histogram contains low values of RSS and BIC, the better the performance of the corresponding model.

distribution of the measured spectrum around the fit, the BIC is given by[26]

$$BIC = 2n \times ln(s) + K \times ln(n).$$

Here, n is the number of data points, s is the standard deviation of the residuals, and K is the number of parameters of the overall model. The BIC includes a penalty term for overly complex models, hence avoiding a preference for models that overfit the data. Fig. 4 shows the distribution of $F_{cost,min}$ and the BIC for all fits over the 10,000 spectra in the database. It can be seen that the generalized Gaussian and the Taylor model generally perform better than the pseudo-Voigt model. The generalized Gaussian model might still perform slightly better than the Taylor model.

2. Representative spectral shapes

The statistical criteria studied above are only one aspect in assessing the fitted model. The fitting model should also conform to the criteria of flexibility, discrimination and robustness, and should be able to capture the salient physics reflected in the spectrum, especially the BB turbulence. In order to validate the better performance of the generalized Gaussian

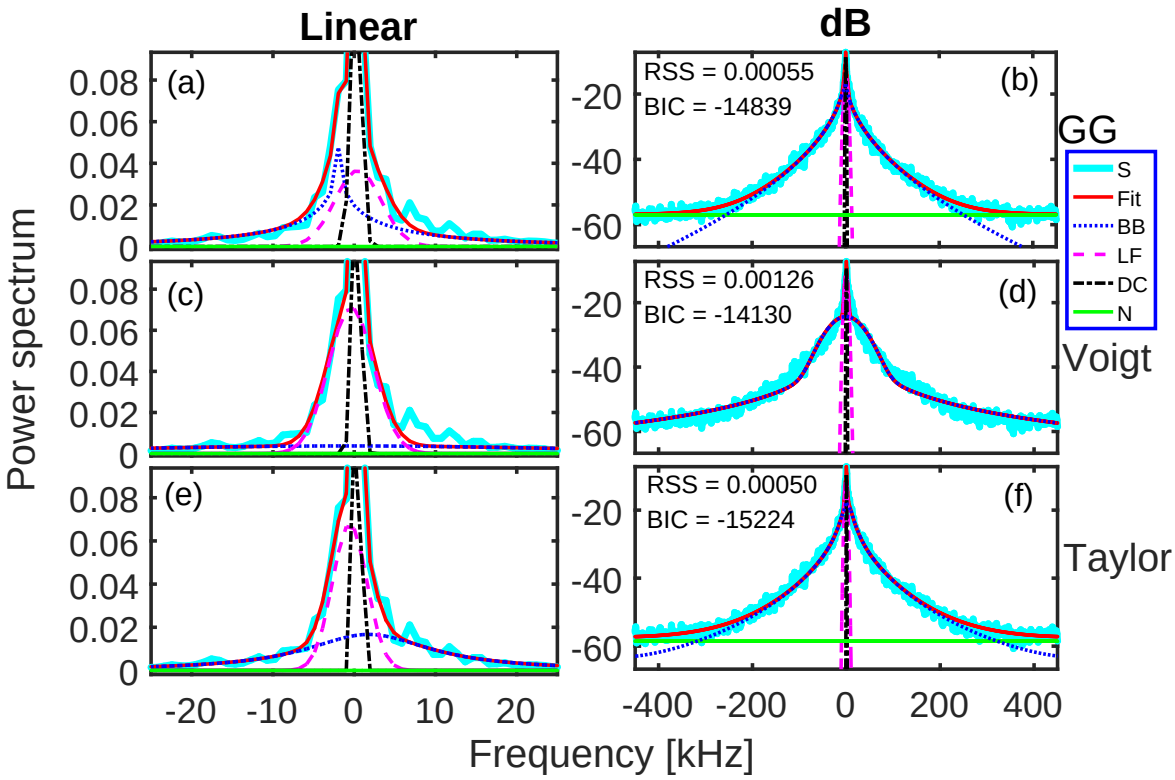

Figure 5. Fit of a Lorentzian spectrum (S), with the individual components also displayed. The BB component was fitted by a generalized Gaussian (GG) function ((a) and (b)), the pseudo-Voigt function ((c) and (d)), and the Taylor function ((e) and (f)). The results are shown on the linear scale ((a), (c), (e)) and logarithmic (dB) scale ((b), (d), (f)). The residual sum of squares (RSS) and the Bayesian information criterion (BIC) at the optimal solution by each model are also displayed.

and Taylor models, as suggested by the statistical analysis, some typical examples were investigated in detail. When the BB component has a Gaussian-like shape, the three models all show an equivalent, excellent fitting performance. In contrast, in case of a more difficult to fit Lorentzian or Laplacian (i.e. double exponential, or triangular on the logarithmic scale), the fitting results can be very different between the three models. This is shown on both the linear and logarithmic scales in Figs. 5 and 6, which correspond to the Lorentzian and Laplacian shape, respectively. On the linear scale, the fit is dominated by the low-frequency part ($f < 25$ kHz), while on the logarithmic scale validation of the fitting performance should concentrate on the larger frequencies (up to 450 kHz).

For the Lorentzian shape in Fig. 5, visual inspection reveals a good fit by all three models, although the fit including the pseudo-Voigt model underpredicts the spectrum between 5 and

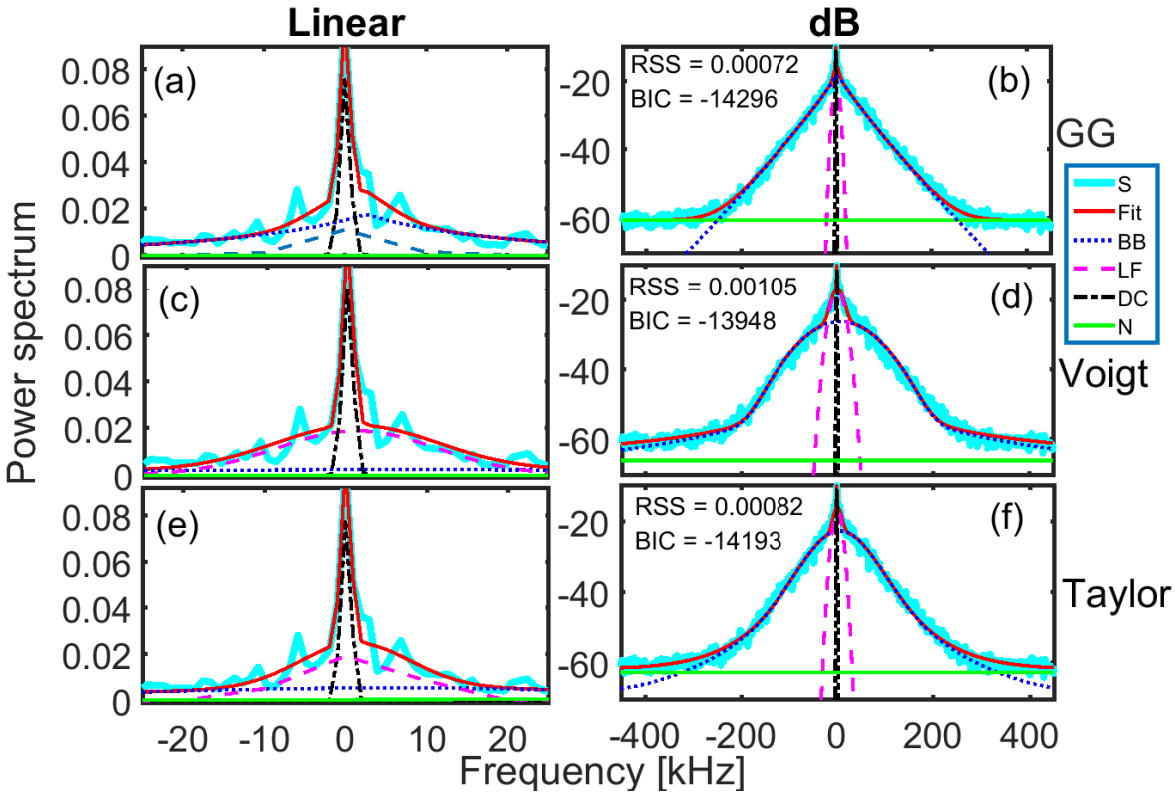

Figure 6. Same as Fig. 5 for a Laplacian shape.

10 kHz (Fig. 5 (c)) and around 100 kHz (Fig. 5 (d)). This is reflected by its slightly worse RSS and BIC compared to the other two models. Another shortcoming of the pseudo-Voigt model is that it tends to fit also the noise, as can be seen in Fig. 5 (d). Similar weaknesses of the pseudo-Voigt function can be seen in Fig. 6. For these reasons, we reject the Voigt function for fitting the BB turbulence.

When comparing the generalized Gaussian (GG) model with the Taylor model, it can be noted in Figs. 5 and 6 that the crucial difference is the peaked shape of the BB component in the GG model, whereas the Taylor model has a much smoother shape. From poloidal correlation in Ref. 15, the BB component disappears when the LF component remains the same, meaning that the BB component does not have an intense low-frequency part. Hence, the fitting results in terms of the BB and LF components can be very different. In Figs. 5 (a) and (e), the average frequencies w.r.t. the central frequency (0 KHz) for the BB and LF components are opposite sign for the two models. Specifically, in Fig. 5 (a), the peaked shape of the GG seems to fit the knee in the spectrum around 3 kHz, whereas this shape is not expected for the BB turbulence. In the case of the Lorentzian or Laplacian spectra,

we observe that the estimated GG shape parameter often saturates at the lower bound $\beta_{BB} = 0.5$, causing a peaked shape that tries to fit small-scale features in the spectrum.

In summary, due to both excellent quantitative and qualitative performance, we choose the Taylor model as the optimal fit to the BB component. Following this choice, various BB turbulence properties, like its energy E_{BB} and spectral shape, can be determined systematically from the parameters of the Taylor model. Nevertheless, the GG model remains useful as a complementary tool for studying the spectral shape, as its parameter β_{BB} has a more straightforward shape interpretation than the parameters of the Taylor model.

IV. APPLICATION TO OHMICALLY HEATED PLASMAS

We now apply our spectrum fitting technique, using the Taylor model for the BB turbulence component, to study Ohmically heated plasmas in Tore Supra. We focused on stationary plasma conditions, removing spectra with low SNR (<25 dB) or large Doppler shift ($\mu_{BB} > 50$ kHz). This resulted in a data set consisting of 180,000 spectra from 3,000 discharges, covering the complete radial extent from the LFS to the HFS. This allowed systematic exploration of the BB energy E_{BB}, obtained from the Taylor fit:

$$E_{BB} = \int_{-500}^{500} C_{BB}^{\mathbf{Taylor}}(f)\,\mathrm{d}f, \tag{11}$$

where the integration is performed over the full frequency range for the BB turbulence. This value lies between 0 and 1, owing to the normalization of the total spectrum.

We have investigated the relation between the turbulence properties obtained from the parametrization and various dimensionless quantities determining the confinement performance. The study was first focused on the radial evolution of the characteristics of the different components of the spectra, with varying edge safety factor. Note that the present study was restricted to the core region and the HFS region ($-1 < \rho < 0.6$), for two reasons. First, when the cut-off layer moves toward the outer plasma edge, most of the spectra are affected by the Doppler effect. This effect can be attributed to the large magnetic field ripple in Tore Supra (> 6% at the edge in standard geometry), distorting the cut-off layers. Second, density fluctuations increase rapidly toward the edge. Indeed, previous studies in Ohmic plasmas have shown that when $\rho \sim 0.6$, the fluctuation level reaches the limit $\delta n/n = 1\%$,[7] a threshold above which non-linear effects can not be neglected. However, it should be

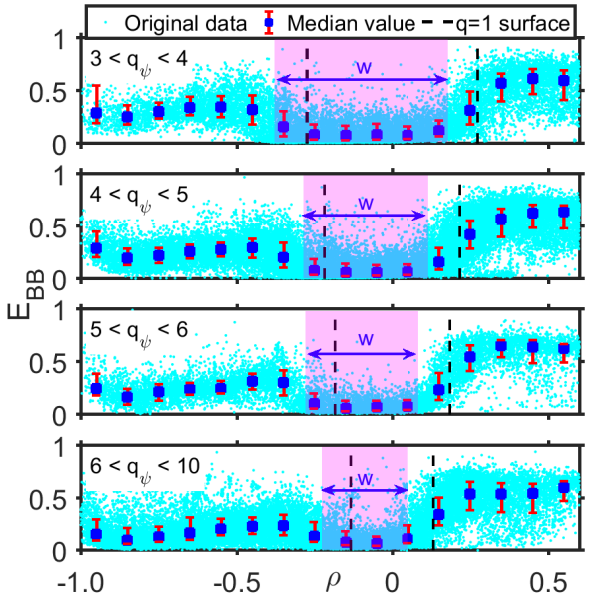

Figure 7. Radial profiles of E_{BB} for different q_ψ as a function of normalized radius $\rho = r/a$, where a is the minor radius of the tokamak ($a \sim 0.72$ m for Tore Supra) and negative values refer to the HFS. The cyan points are obtained from the individual fitted spectra. The deep blue square points denote median values calculated within small radial intervals ($\delta\rho \sim 0.1$), with red error bars around the median given by the mean absolute deviation. The $q = 1$ positions are indicated by the black dashed lines.

noted that the effect of the high-level edge turbulence on the core measurements is limited, as the waves that are multiply scattered by the edge fluctuations, potentially blurring the information on the core turbulence, are spread out and are often not detected.[27,28]

A. Dependence of E_{BB} profile on q_ψ

Figure 7 shows the radial profiles of the broadband energy E_{BB} content of the normalized spectra for different ranges of the edge safety factor q_ψ. The most remarkable feature is a clear reduction of E_{BB}, named the *energy basin*, in the core region for all ranges of q_ψ. Furthermore, there is a clear asymmetry between the HFS and LFS: E_{BB} tends to slowly increase from the inner edge towards the center up to the cliff before the energy basin on the HFS, whereas on the LFS E_{BB} is much higher and reaches saturation level above $E_{BB} > 0.5$, indicating that the BB component prevails in the reflected spectra.

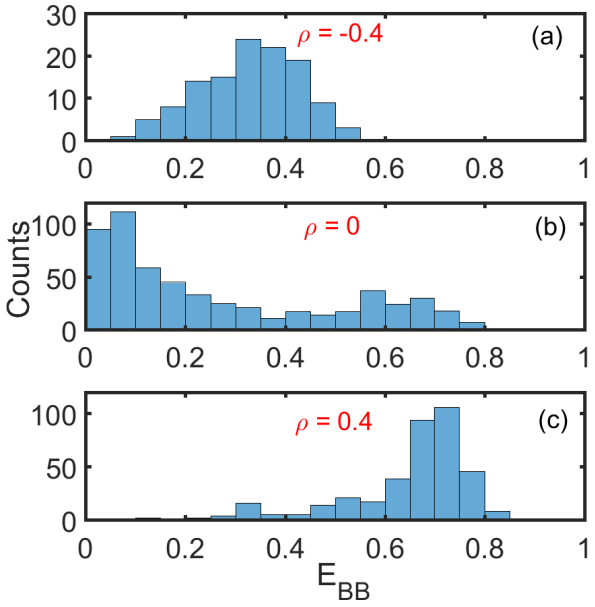

Figure 8. Different distributions of E_{BB} at fixed radial positions: (a) $\rho = -0.4$, (b) $\rho = 0$, (c) $\rho = 0.4$, under the condition $5 < q_\psi < 6$.

Specifically, the radial evolution is not due to the distinct averaging of the different components of the reflectometry spectra caused by the beam size effect. The beam size is expected to be larger than the BB turbulence wavelength. For the LF component, its wavelength could be longer than the beam size at the edge, while it can become comparable to the beam size due to the decreasing minor radius close to the center. Unfortunately, Doppler reflectometry, which can evaluate the poloidal correlation length, is not available for core and HFS measurements. However, if such a beam size effect were predominant, one would neither observe a modification of the position of the energy basin with varying edge q profile, nor the rise of E_{BB} on the HFS from the edge up to the basin cliff.

The drop of the density fluctuation levels inside the q=1 surface was already reported[7] but for a limited number of discharges and plasma parameters. The database analysis shows that this drop is a generic observation in Ohmic cases. To go further, a link between the BB energy content and the density fluctuation level should be established. Correction factors like the wavelength of the probing wave, the scale length of permittivity and the radial correlation length of the turbulent structures should be taken into account[3,20]. To establish a link between the normalized BB component and the fluctuations level, the differences between the LF and BB radial correlation lengths should also be considered.

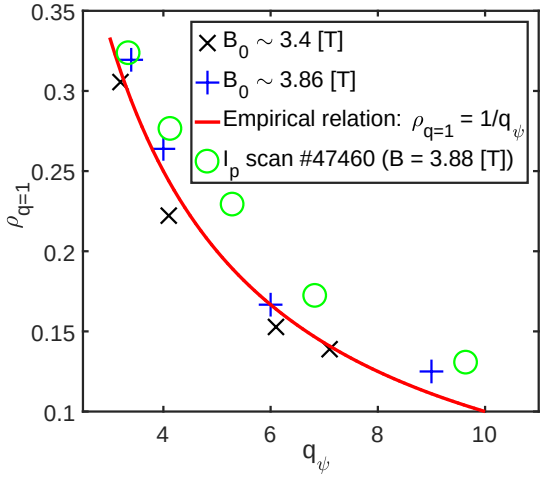

Figure 9. Verification of the empirical relation $\rho_{q=1} = 1/q_\psi$ by some typical discharges from the Tore Supra database.

On the other hand, Fig. 7 shows that, at fixed radial position, E_{BB} still varies considerably across the database, for all values of q_ψ. This can be attributed to fitting errors and varying global operating conditions. In Fig. 8, the distributions of E_{BB} are shown at three radial positions ($\rho = -0.4, 0, 0.4$), in the range $5 < q_\psi < 6$, where the variance of E_{BB} is the lowest. Apart from the large scatter of E_{BB} at fixed radial position, differences in the mean and shape of the distributions are apparent. Because of the non-zero skewness and outliers in the distributions, we use the median of the distribution rather than the mean for systematic studies of the typical broadband energy. When calculating this within small radial intervals of width ~ 0.1, we obtain the deep blue squares in Fig. 7. The red error bars are given by the mean absolute deviation around the median values within each interval.

B. Relationship between E_{BB} basin and $q = 1$ surface

In Fig. 7 it can also be seen that, as q_ψ increases, the energy basis shrinks. This suggests a relation between the E_{BB} basin and the $q = 1$ surface. However, at Tore Supra reconstruction of the $q = 1$ surface from the routine equilibrium reconstruction is affected by considerable uncertainties. Therefore, the position was estimated through the approximate empirical relation $\rho_{q=1} = 1/q_\psi$, established earlier for TFTR and TFR.[29]. To verify this relation for our database, we employed several typical discharges in different toroidal magnetic field B_0

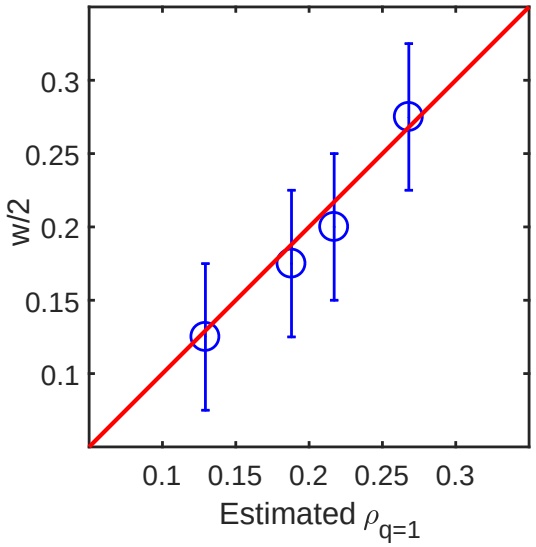

Figure 10. Half-width of E_{BB} basin as a function of the q=1 position, both values normalized to the minor radius.

(3.4 T and 3.86 T), as well as a pulse with an I_p scan. The result is shown in Fig. 9, confirming the validity of this empirical relation and providing a practical means to derive the position of the $q = 1$ surface. In each q_ψ range, the median value using $\rho = 1/q_\psi$ gives an approximation to the position of the $q = 1$ surface, shown as the two vertical dashed lines in Fig. 7.

For a quantitative definition of the width of the E_{BB}, we employ the criterion $E_{BB} < 0.1$ (value of the median). The radial region of the basin and its width (w) are indicated by the shaded area and the double arrow shown in Fig. 7. In addition, the half-width of the basin (w/2) is shown as a function of the normalized q=1 position ($\rho_{q=1}$) in Fig. 10. There is a clear one-to-one correspondence, supporting our hypothesis that the occurrence of the E_{BB} basin is related to the $q = 1$ surface. The error bars on the half-width originate from the limitation on the spatial resolution due to the finite number of points in each radial interval. The transition of broadband energy level across the $q = 1$ surface requires further study.

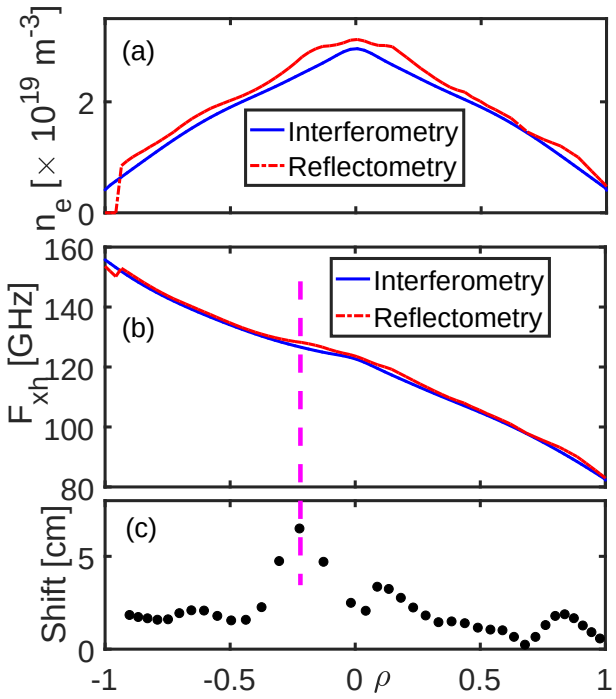

Figure 11. (a) Density profiles and (b) upper cutoff frequencies (F_{xh}) near the turbulence signal obtained by core reflectometry and interferometry. (c) Difference of the cutoff positions from interferometry w.r.t. reflectometry at different radial positions.

C. Shift of the cutoff layers

Although the half-width of the E_{BB} basin correlates well with $\rho_{q=1}$, Fig. 7 reveals that the actual position of the boundaries of the basin do not coincide with the location of the $q = 1$ surface, indicated by the shaded basin region. A systematic shift towards the HFS of the radial positions is observed in Fig. 7. To understand the origin of this shift, it is important to recall that the radial positions are the cutoff positions calculated using the density profile from the interferometry diagnostic. Hence, the shift might be due to uncertainties on the density profiles from interferometry. To resolve the matter, we investigated several tens of Ohmic discharges with available core reflectometry profiles. Usually, interferometry underestimated the core density compared to the reflectometry. This might be due to the density plateau and complex profile structures near the magnetic axis.[7] To illustrate this difference Fig. 11 (a) shows the density profiles from core reflectometry and interferometry, acquired at the same time in one typical discharge. It can be seen that the core density from interferometry is lower than the reflectometry one. We next calculated the cutoff positions

using each of the two density profiles, confirming a shift towards the HFS of the cutoff layers from interferometry w.r.t. reflectometry, as plotted in Fig. 11(c). There is a clear asymmetry, with a much larger shift in the region $-0.3 \lesssim \rho \lesssim -0.1$ compared to $0.1 \lesssim \rho \lesssim 0.3$, i.e. in the vicinity of the $q = 1$ surface. This is consistent with the observation in Fig. 7, where the $q = 1$ surface is outside the E_{BB} basin on the LFS, but inside the basin on the HFS, with an asymmetry due to the larger shift. The strong shift in the HFS region is caused by the slower increase of the cutoff frequencies (Fig. 11 (b)) in the HFS than in the LFS.[3] The change of shift is relatively large even with only a small change of the probing frequencies, since the evolution of X-mode upper cutoff frequency with ρ becomes flatter in the HFS, due to the fact that the magnetic field intensity continues to increase, while the density gradient changes sign across the magnetic axis. Thus, the peak of the shift in Fig. 11 (c) corresponds to the flattest part of the cutoff frequency profile deduced from the reflectometry density profile, as indicated by the vertical dashed line in Fig. 11 (b) and (c).

V. CONCLUSIONS AND OUTLOOK

A robust spectrum parametrization method for systematic studies of turbulence properties in fusion plasmas has been developed, as a complementary tool to the traditional analysis method on the basis of a limited number of key discharges. This is intended to open the way to a new, standardized method for studying plasma density fluctuations (δn) from a systematic viewpoint, in order to reveal global patterns or trends that are difficult to find by studying only a limited set of plasma conditions. We have described a number of steps to derive the parametrization in a comprehensive database of fluctuation reflectometry data from Tore Supra plasmas. The method is also useful for quantifying the power spectrum in individual discharges, and can be easily adapted to other fusion devices or other research domains relying on quantitative comparison of spectra.

The generalized Gaussian, Voigt, and Taylor models have been used to parameterize the fluctuation power spectrum. The Taylor model gives the best performance in terms of goodness-of-fit and BIC, while meeting the requirements of flexibility, discrimination, and robustness. In implementing the fitting routine, the cost function, the constraints and the initial guesses have been identified as critical points. The cost function has the same weight for the linear and logarithmic scale, while the constraints are used for separating the differ-

ent components in the spectrum and multiple initial guesses guarantee global convergence. However, in this work some typical values were chosen empirically, which could be different when adding more components in the model or when applying the method to another database.

In this paper a first demonstration of a systematic study of a particular feature of the fluctuation power spectrum has been given in Ohmically heated Tore Supra plasmas under stationary plasma conditions. The complete radial profile of the broadband energy E_{BB} has been investigated for different edge safety factors q_ψ. An E_{BB} basin in the core region was shown to reflect the position of the $q = 1$ surface, and its in-depth study is underway. The systematic shift of the cutoff layers to the HFS is probably caused by underestimation of the electron density in the core region by the interferometry diagnostic.

In future work, we will use the database we have built with this method to investigate the origin of the difference of spectral shape in the LFS and HFS. We want also to establish a link between the BB component and the density fluctuation level for more in-depth studies of plasma turbulence properties, comparing the linear Ohmic confinement (LOC) and saturated Ohmic confinement (SOC) regimes, investigating the effect of additional heating schemes, etc.

ACKNOWLEDGMENTS

One of the authors (Y. Sun) wishes to thank L. Vermare and P. Hennequin for instructive discussions about spectrum fitting by the Taylor model and D. Elbeze for the analysis tool which was helpful in validating the position of the sawtooth inversion. Furthermore, the authors sincerely appreciate the reviewers' comments which lead us to carry out a deeper investigation of the observations and to improve the construction of the manuscript. This work has been carried out within the framework of the Erasmus Mundus International Doctoral College in Fusion Science and Engineering (FUSION-DC).

REFERENCES

[1]N. Bretz, Rev. Sci. Instrum. **68**, 2927 (1997).

[2]R. Nazikian, G. J. Kramer, and E. Valeo, Phys. of Plasmas **8**, 1840 (2001).

[3]E. Mazzucato, Rev. Sci. Instrum. **69**, 2201 (1998).

[4]G. D. Conway, Plasma Phys. Control. Fusion **39**, 407 (1997).

[5]G. D. Conway, Plasma Phys. Control. Fusion **50**, 124026 (2008).

[6]T. Gerbaud, F. Clairet, R. Sabot, A. Sirinelli, S. Heuraux, G. Leclert, and L. Vermare, Rev. Sci. Instrum. **77**, 10E928 (2006).

[7]R. Sabot *et al.*, Plasma Phys. Control. Fusion **48**, B421 (2006).

[8]G. D. Conway, B. Scott, J. Schirmer, M. Reich, A. Kendl, and the ASDEX Upgrade Team, Plasma Phys. Control. Fusion **47**, 1165 (2005).

[9]N. Crocker *et al.*, Phys. of Plasmas **15**, 102502 (2008).

[10]M. V. Ossipenko and the T-10 team, Nuclear Fusion **43**, 1641 (2003).

[11]A. Krämer-Flecken, V. Dreval, S. Soldatov, A. Rogister, V. Vershkov, and the TEXTOR-team, Nucl. Fusion **44**, 1143 (2004).

[12]V. A. Vershkov, D. A. Shelukhin, S. V. Soldatov, A. O. Urazbaev, S. A. Grashin, L. G. Eliseev, A. V. Melnikov, and the T-10 team, Nucl. Fusion **45**, S203 (2005).

[13]H. Arnichand *et al.*, Nucl. Fusion **54**, 123017 (2014).

[14]V. A. Vershkov *et al.*, Nuclear Fusion **51**, 094019 (2011).

[15]A. Krämer-Flecken, S. Soldatov, Y. Xu, H. Arnichand, S. Hacquin, R. Sabot, and the TEXTOR team, New Journal of Physics **17**, 073007 (2015).

[16]R. Sabot, A. Sirinelli, J.-M. Chareau, and J.-C. Giaccalone, Nucl. Fusion **46**, S685 (2006).

[17]R. Sabot, C. Bottereau, J.-M. Chareau, F. Clairet, and M. Paume, Rev. Sci. Instrum. **75**, 2656 (2004).

[18]C. H. dos Santos Amador, R. Sabot, X. Garbet, Z. G. aes Filho, and J.-H. Ahn, Nucl. Fusion **58**, 016010 (2017).

[19]P. Welch, IEEE Transactions on Audio and Electroacoustics **15**, 70 (1967).

[20]D. A. Shelukhin, S. V. Soldatov, V. A. Vershkov, and A. O. Urazbaev, Plasma Physics Reports **32**, 707 (2006).

[21]T. Ida, M. Ando, and H. Toraya, J. Appl. Cryst. **33**, 1311 (2000).

[22]P. Hennequin, C. Honoré, A. Quéméneur, A. Truc, F. Gervais, C. Fenzi, and R. Sabot, in *26th EPS Conf. on Contr. Fusion and Plasma*, Vol. 23J (1999) pp. 977 – 980.

[23]D. Grésillon *et al.*, Plasma Phys. Control. Fusion **34**, 1985 (1992).

[24]C. Gil *et al.*, Fusion Science and Technology **56**, 1219 (2009).

[25]P. Thompson, D. E. Cox, and J. B. Hastings, J. Appl. Cryst. **20**, 79 (1987).

[26]W. von der Linden, V. Dose, and U. von Toussaint, *Bayesian Probability Theory, Application in the Physical Sciences* (Cambridge University Press, 2014).

[27]G. Zadvitskiy, *Turbulence studies in fusion plasmas using reflectometry stnthetic disgnostic*, Ph.D. dissertation, University of Lorraine (2018).

[28]E Z Gusakov and A Yu Popov, Plasma Phys. Control. Fusion **44**, 2327 (2002).

[29]V. Arunasalam *et al.*, Nucl. Fusion **30**, 2111 (1990).