EXTRACTION OF QUASI-COHERENT MODES BASED ON REFLECTOMETRY DATA

Luigui Salazar[1,2], Stéphane Heuraux[1], Roland Sabot[2] and Andreas Krämer-Flecken[3]

[1] IJL UMR 7198 CNRS, Université de Lorraine, F-54000 Nancy, France
[2] CEA, IRFM, F-13108 Saint-Paul-Lez-Durance, France
[3] Institut für Energie- und Klimaforschung / Plasmaphysik, Forschungszentrum Jülich GmbH, D–52425 Jülich, Germany

E-mail: luigui.salazar@cea.fr

Abstract.

The identification of turbulence sources would drive to a deeper understanding of confinement dynamics in tokamak plasmas. Turbulence results from a mixture of instabilities corresponding to sources at different time and spatial scales. Using correlation reflectometry and multipin Langmuir probe (MLP), it was shown in the T-10 and TEXTOR tokamaks that the reflectometry frequency spectrum was the superposition of several components: broadband component, Quasi-Coherent (QC) modes and low frequency components. The relevance of QC modes is associated to their link with the Trapped Electron Mode (TEM) instability. A method is presented to extract the QC modes component from the reflectometry data enabling the separation from the broadband component and the study of its time evolution. It is a first step toward the discrimination of turbulence sources. The central idea explores a way to combine the approach of signal processing and machine learning: the Continuous Wavelet Transform (CWT) on the basis of complex Morlet wavelet has proven to be efficient in providing a decomposition of a signal at different scales over time for fluctuations tackling; clustering techniques as the Minibatch K-means are able to tackle clusters at different scales. The method was applied to Tore Supra and TEXTOR reflectometry data. On Tore Supra, the amplitude of the extracted QC mode component decreases during transition from the Linear Ohmic Confinement (LOC) to the Saturated Ohmic Confinement (SOC) regimes. On TEXTOR the amplitude of the coherent spectra of the extracted QC mode component is similar to the experimental coherent spectra obtained through correlation reflectometry. The developed method permits to extract components preserving the physical contain and the statistical properties of it.

plasma turbulence, reflectometry, diagnostic, tokamak

1. Introduction

Turbulence is the root of confinement degradation in tokamaks since it induces an anomalous transport of energy and particles [1]. It is strongly non-linear and it is the result of instabilities such as the Ion Temperature Gradient (ITG), the Trapped

Electron Mode (TEM) and the Electron Temperature Gradient (ETG). The sources originating these instabilities are believed to reign turbulence transport at different regimes for time and spatial scales. Consequently, tackling these sources would lead to a better understanding of dynamics among instabilities. According to harmonic analysis, they are the superposition of various spectra at several frequency-bands [2]. It is generally assumed that there is no coupling between these various instabilities [3], but an experimental proof would be useful.

Among the different diagnostics able to probe turbulence and fluctuations of the plasma parameters, reflectometry, a radar-like technique based on the propagation of electromagnetic waves in plasmas, has been found very versatile [4] for the analysis of density fluctuations as it requires simple access to the plasma and provides a good time and space resolution.

For the purpose of investigating turbulence properties, a study was carried out using correlation reflectometry (CR), multiplin Langmuir probe (MLP) and heavy ion beam probe (HIBP) in the T-10 tokamak [5]. This work succeeded in splitting the frequency spectra in several components: broadband, low and high frequency Quasi-Coherent (QC) modes, low frequency components and some oscillations at 15-30 kHz. Similar observations were reported in the TEXTOR tokamak [6]. In both tokamaks, the amplitude of the high frequency QC mode is much lower than the low frequency one.

It is relevant to have a deeper look at the low frequency QC modes. It was shown that they are a signature of the TEM instability [7]. Indeed, a disappearance of these modes was reported in several tokamaks [8, 9] during transitions from the Linear Ohmic Confinment (LOC) to the Saturated Ohmic Confinement (SOC) regimes. The SOC-LOC transition is attributed to the stabilization of the TEM [10]. However, results from [11] were based on an average behaviour, de facto excluding the short time scale evolution of the frequency spectra.

A complementary study has been performed on TEXTOR. The QC-modes exhibit short and long range correlation as shown by the significant values of the magnitude squared coherence between antennas at long poloidal and toroidal distances [6, 12]. Long time series permit to observe this long range coherence and in the following offer the possibility to investigate the short time scale behavior.

These first investigations on spectrum decomposition were continued at the Tore Supra tokamak. The aim was to automatically decompose the reflectometry frequency spectrum into a set of four components [13] -direct current component, low-frequency fluctuations, broadband turbulence and the noise level- to perform systematic analysis. A database of 6000 discharges database and more than 350 000 reflectometry measurements was built to extract trends from statistical analysis instead of dedicated shot-to-shot analysis [14]. Nevertheless, this decomposition included some limitations as the assumption of a stationary turbulence behaviour during the reflectometry probing time. Moreover, side-lobe components such as QC-mode and high frequency MHD modes were integrated in the broadband component.

From this point of view, the long term objective is to overcome the stationary

assumption and thereby offer a way to tackle the dynamics of the turbulence components. The Continuous Wavelet Transform (CWT) was used to capture density fluctuations properties from the reflectometry raw complex signals at various scales [15].

The extraction of the low frequency QC mode component is based on a screening technique applied to the CWT decomposition. The screening is composed of two threshold functions: cross-correlation threshold and density energy threshold. An union of the patterns of both screening domain is chosen to combine correlation and energy density properties. However, this choice can include irrelevant elements for QC mode characterization.

To finish the QC-mode extraction, we rely on Artificial intelligence (AI). AI is an area that is evolving fast during the last years. The first step of the IA procedure is to find a vector space for merging properties of CWT similar to the density energy and wavelet scales. The second AI step is the Mini-batch K-means (unsupervised clustering technique) to tackle the cluster corresponding to the QC modes.

More "classical" methods such as short time Fourier transform, PCA (principal component analysis) and variational auto-encoder network were first explored but they did not succeed: prohibitive calculation time for an application to a large database and low signal to noise ratio (SNR).

Our method has proved successfully applied to Tore Supra and TEXTOR reflectometry data for extracting the QC mode components. The decrease of the QC-mode amplitude during LOC-SOC transition in Tore Supra is recovered. On TEXTOR, the magnitude squared coherence (MSC) obtained from two poloidaly separated antennas is very similar to the coherence obtained after applying the QC mode extraction.

The extraction method is thus able to eliminate other components from the spectrum without affecting the QC component. In this sense, a better decomposition of reflectometry data spectrum is reachable by removing the QC modes from the broadband component, offering a simpler signal to analyze for further studies.

2. Exploratory part and final data analysis method

2.1. General approach

Before finalizing the method, different attempts and approaches were tested to extract the spectrum components. Since the AI algorithms are so numerous and diverse, we think it is worth to shortly present the ones we tried and their shortcomings. Indeed, application of AI to reflectometry and more generally fusion plasma turbulence is still ongoing and thus the most suitable methods have not yet emerged.

2.1.1. Metric DBSCAN First a DBSCAN [16] was tested for finding the QC modes directly from the raw complex signal. DBSCAN is a technique to identify clusters in a multidimensional space. To find a space where the different components would be

separated, we tried the general relativity metric concept in which the tensor field varies among elements [17]: the central idea was to find invariant values by optimizing α and β coefficients (1) . After this process, DBSCAN was applied in order to seek for density clusters over the metric representation.

$$Signal : A_{(t)}.e^{\phi_{(t)}}, dS^2 = \alpha_{(\phi)}.dA^2 + \beta_{(t)}.dt^2 \tag{1}$$

Although some structures emerged, it was difficult to identify the modes, with the main issue associated to the computing time and the huge quantity of data. We didn't pursue this path at this first step.

2.1.2. Statistical methods This statistical approach belongs to the classical methods in which the idea is to smooth the raw signal spectrum to improve its shape and reduce the noise. The moving average filter and the local average filter were able to achieve some smoothing, but the uncertainty was too high due to the lack of information on the shape of the different spectral components.

To overcome these issues, a time-frequency representation has the potential to separate the different components by their dynamics. Besides, it allows to access to the short time evolution.

2.2. Time-frequency representation

2.2.1. Short time Fourier transform The natural approach for Time-Frequency Representation (TFR) is the Short Time Fourier Transform (STFT). STFT gives a 2D representation (image) on which convolutional filters can be applied.

$$\begin{aligned} &Sf(u,\varepsilon) = \langle f, g_{u,\varepsilon} \rangle = \int_{-\infty}^{+\infty} f(t)g(t-u)e^{-i\varepsilon t}dt \\ &u : time\,step \\ &\varepsilon : frequency \end{aligned} \tag{2}$$

Nevertheless, the Heisenberg box associated to this transform is constant along frequency and time axis resulting a mixing of time and frequency scales due to the convolution between the time window FFT and the observed frequencies leading to a blurred TFR. Wavelet transform [15] provides a variable Heisenberg box in TFR.

2.2.2. Continuous wavelet transform Continuous Wavelet Transform (CWT) [18] was chosen rather than the Discrete Wavelet Transform (DWT) for the following reasons: it permits to analyze time series that contain a non stationary power at many different frequencies [19] -the scales can be adjusted to those of the observed fluctuations- whereas the scales of the DWT are limited to a discrete set [20].

Some parameters are essential in the definition of this transform: the type of wavelet and the wavelet dictionary i.e all wavelet versions which are going to be used [15]. This transform is based on a non-orthogonal analysis.

$$\begin{aligned} &Wf(u,s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^*(\frac{t-u}{s}) dt \\ &u : time\, step \\ &s : scale \end{aligned} \tag{3}$$

- Wavelet :

As reflectometry signals are complex, a complex wavelet must be used in order to tackle angle (position) and module (energy). Two wavelets were brought out for this operation: Paul wavelet and Morlet wavelet [21]. Paul wavelet revealed itself to be efficient for phase detection whereas Morlet wavelet is prioritized for detection of fluctuations, our main goal. The Morlet wavelet is defined as:

$$\begin{aligned} &\psi_{morlet}(t) = |s|^{-0.5} \cdot \pi^{-0.25} \cdot e^{6(\frac{t \cdot f_s}{s})j} \cdot e^{-0.5(\frac{t \cdot f_s}{s})^2} \\ &s : scale \\ &f_s : frequency\, sampling \end{aligned} \tag{4}$$

- Dictionary:

The wavelet dictionary is created from scales and time steps, in this whole process the time step is set to be the sampling period ($u = \delta(t)$), then scales are defined in such a way that the calculation time is minimized and the frequency band of concern is covered ([20kHz$-f_s/2$]), removing the low frequency components ($[0 - 20KHz]$).

$$D = \{\psi_{u,s(t)} = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right) \} \tag{5}$$

It should be pointed out that negative and positive scales are presented because of complex raw signals. Negative scales correspond to the negative frequency bandwidth, likewise for positive scales.

The first approach on CWT was to take every scale and to perform Independent component Analysis (ICA) at every time step in order to separate the CWT spectrum into additive sub-components by assuming that they are non-Gaussian signals. Main constraints were the lack of precision for reconstruction trough the Inverse Continuous Wavelet Transform (ICWT) of CWT to obtain QC modes.

2.2.3. Other approach It relies on a variational auto-encoder network for performing encoding of CWT into a multivariate latent distribution and decoding for reconstruction along the time axis, the idea was to divide CWT into several windows in order to obtain a pattern window evolution. Main constraints were calculation time and learning process.

The CWT approach proved to be suitable to highlight patterns in TFR, but additional algorithms tackling different properties of the QC modes were needed to extract them.

2.3. Final approach and Data analysis procedure

Finally, in this section a description of the algorithm process from the reflectometry database to the extraction of QC modes is done. The High Frequency HF-QC modes, when detected, are observed at frequencies close to the first harmonic of the low frequency QC modes ones making the identification of HF-QC modes difficult. For the rest of the paper, the QC mode term refers to the low frequency QC mode, ie $f < 150$kHz.

Complex raw signals ($f(t)$) are selected (dataframe) based on pre-defined plasma conditions, and undergoes a CWT, a screening, a clustering and finally an ICWT. The Fig. 1 illustrates the procedure applied to a reflectometry aquisition (#32093 $t = [11.1151s, 11.1397s]$, $F = 118.8$ GHz) that exhibits strong QC modes.

2.3.1. Screening technique: Once the CWT is obtained (Fig 2), the chosen reference scale is the one for which the local average *Energy Density*(ED) is maximum (s_{hed}) (see Fig 2) in the frequency range associated to the QC mode. It is done separately for negative and positive scales.

For the screening, hard threshold techniques are used. The threshold are estimated from reflectometry measurements with strong QC amplitude, thus the QC-modes are more likely to be discriminated. Those signals $g_n(t)$ belong to the Tore Supra shot range of #32263 to #32265.

- Cross-correlation threshold:

The minimum of *Cross-Correlation*(CC) is used as a threshold to enhance the Signal over Noise ratio (SNR), the latter can be deduced from (6):

$$SNR = \frac{\mu(|Wg_n(u,s)|)}{\sigma(|Wg_n(u,s)|)} \tag{6}$$

where $Wg_n(u,s)$ is the CWT of the chosen signal $g_n(t)$.

The threshold is then estimated :

$$Threshold_{CC} \approx \frac{\sum_{n=1}^{N} SNR_n}{N} \tag{7}$$

Then, a cross-correlation operation is performed between s_{hed} and others scales s (either positive or negative scales) and then normalized by their standard deviation as shown by (8).

$$CC_{normalized} = \frac{|s_{hed}(t)|^2 \otimes |s(t)|^2}{\sigma_{hed} \cdot \sigma} \tag{8}$$

Finally, hard thresholding (T_{h1}) is applied (Fig 3):

$$T_{h1} = \begin{cases} Wf(u,s), & |CC_{normalized}| \geq Threshold_{CC} \\ 0, & |CC_{normalized}| < Threshold_{CC} \end{cases} \tag{9}$$

- Energy density threshold:

This threshold is based on the minimum Energy Density for existence of QC modes. In that regard, we take every $|Wg_n(u,s)|^2$ from all u and s and project it into a 1-dimensional vector. At the beginning, we tried to estimate its probability density function (PDF) by doing convolutions with pre-defined kernels as Gaussian and Cosine. Nevertheless, it was not efficient because of low precision in distinguish components. This led us to a parametric approach in which we could pre-define the number of components or clusters.

According to Yan Sun's spectrum decomposition [13], once the two low frequency components (direct current and low frequency fluctuations) are filtered out, two components remain: the broadband turbulence and the noise. Since we considerate QC modes as a third component, we choose 3 clusters to apply the K-means algorithm [22], an unsupervised machine learning algorithm that minimize the within-cluster variances. Consequently, as there are 3 clusters, 2 limits are defined among them.

$$\begin{aligned} ED_{(cluster1)} &\in [min1, max1\rangle; \\ ED_{(cluster2)} &\in [max1, max2\rangle; \\ ED_{(cluster3)} &\in [max2, max3] \end{aligned} \tag{10}$$

Then,

$$Threshold_{ED} \approx max1 \tag{11}$$

Finally, hard thresholding (T_{h2}) is applied (Fig 4):

$$T_{h2} = \begin{cases} Wf(u,s), & |Wf(u,s)|^2 \geq Threshold_{ED} \\ 0, & |Wf(u,s)|^2 < Threshold_{ED} \end{cases} \tag{12}$$

The objective is to perform a mesh able to catch at least QC modes and thereby eliminate most of the noise. To achieve this objective, the ED and CC "patterns" resulting from the hard thresholding undergo an union operation in order to combine properties of the ED and CC thresholds.

$$Wf(u,s)_{screened} = T_{h1} \cup T_{h2} \tag{13}$$

2.3.2. Clustering technique A second process is needed for extracting the QC modes from the screened CWT representation.

- Energy density-scale vector space:

The choice of this vector space relies on combining the properties from ED and CWT scales (named s) identified by the screening technique. This vector space follows the characteristics of an Euclidean space $\mathbb{R}^2$. First, all $|Wf(u,s)|^2$ and s belonging to $Wf(u,s)_{screened}$ are normalized as following:

$$\begin{aligned} s_{normalized} &= \frac{s}{s_{max}} \\ |Wf(u,s)|^2_{normalized} &= \frac{|Wf(u,s)|^2}{|Wf(u,s)|^2_{max}} \end{aligned} \tag{14}$$

Then, the vector space is set as following:

$$\mathbf{V} = (s_{normalized}, |Wf(u,s)|^2_{normalized}) \tag{15}$$

And, it's represented as (Fig 5):

- Clustering:

This part seeks to identify the QC mode cluster in the vector space $\mathbf{V}$. It is based on Minibatch K-means: compared to the K-means algorithm, it's main idea is to use small random batches of examples of a fixed size so they can be stored in memory, in each iteration a new random sample from the dataset is obtained and used to update the clusters and this is repeated until convergence [23].

As discussed in the section *2.3.1*, the number of components or clusters existing are equal to 3, besides some other parameters has to be set for the minibatch K-means algorithm: k-means++ method[24] for initialization of cluster centers, batch size equal to 1024 and the random state for random number generator equal to 0. This algorithm is thus applied into the Energy density-scale vector space $\mathbf{V}$ (Fig 6).

QC modes cluster are identified thanks to the s_{hed} chosen previously - this scale has to be present in the QC cluster.

As mentioned previously, this procedure is performed separately for negative and positive scales. Once QC mode cluster is identified for both of them, they are back projected into CWT space for reconstruction, ICWT is applied to obtain QC mode cluster in time representation (Fig 7).

Thus, we can obtain the frequency spectrum of the initial signal, QC modes and the initial signal without QC modes(Fig 8):

3. Application of the algorithm

This part describes the application of our method to the experimental data from Tore Supra and TEXTOR reflectometers.

In Tore Supra, the reflectometry system worked in the frequency range of 105-155 GHz in X-mode [25]. The acquisition was performed at 1 MHz sample frequency.

Typically series of 20 to 30 frequency steps were performed several times per shot. The frequency step acquisition time windows was short about 10 ms to 20 ms.

In TEXTOR, the reflectometer worked in O-mode in the frequency range of 26-40 GHz. An antenna system built with one emitter and two receivers antennas installed symmetrically to the emitter in the same cross section enabling to perform short range poloidal correlation. The sampling frequency was 500KHz and collected the data for a time interval of $2s$ within a TEXTOR discharge [6].

3.1. LOC-SOC transition

The LOC-SOC transition was first identified by the saturation of the energy confinement time with increasing plasma density, see Rice's review [10]. The transition is associated to the change of the dominant turbulent instability. Linear stability analysis performed with the GYRO code shows that TEM instability are dominant for long wavelength turbulence in the LOC regime and ITG modes are dominant in the SOC regime [26]. Recent results from gyro-kinetic codes reach similar conclusions [27]. Different kinds of signature are also associated to this behaviour change: the flip of the turbulent phase velocity [28, 29], change of the plasma toroidal velocity direction [30], the rollover in density peaking occurs [31], change in the electron fluctuation frequency spectra [26]. Finally, to obtain an estimation of the LOC-SOC transition (Fig 9) instead of analyzing a current ramp-up it is recommended to analyze a density step-by-step scan [10].

An important fact associated to the LOC-SOC transition is the disappearance of the QC modes in the reflectometry frequency spectra [8, 9]. Instead of showing a broadband spectra as in the LOC regime, the density fluctuation spectra exhibit a peak of the order of 10 kHz in the 20-100 kHz range. The bandwidth is intermediate between narrow coherent modes and the broadband turbulence component. It was recently shown that this bandwidth can evolve from a coherent peak to a broadband turbulence with increasing collisionnality in ECRH plasmas [32].

In these previous studies, no quantification was performed. We apply our method to the the same shot range as in [33] in order to quantify the energy contribution of the QC mode to the frequency spectrum. Fig 10 shows the extracted QC modes for the same acquisition as those shown in figure 1c from [33]. The method could not be applied successfully to the last spectrum $n_e = 4.6 \cdot 10^{19} m^{-3}$ because the spectrum is affected by Doppler effects, as a consequence, the method detects a QC mode with the wrong amplitude (also in d and f). The main part of the QC mode contribution is extracted. An harmonic or an HF QC mode is present (a and b) and disturb the extraction.

As the QC-mode extraction process is quite fast (30 seconds for 10 000 time points), it was applied to all the reflectometry measurements in the LOC-SOC shot range for which the cut-off layer was at $r/a = 0.3 \pm 0.025$. After excluding data with low SNR and high Doppler shift, the QC mode energy normalized to the Broadband energy shows a clear decrease of the QC mode contribution with increasing density (Fig 11). The transition is around a line average density $< n_l >= 3 \cdot 10^{19} m^{-3}$. This threshold is

similar to the density threshold for the time confinement observed in the Fig 9.

To verify that the signal properties are conserved after applying our method, we have a look at the coherence conservation between two antennas of poloidal correlation reflectometers developed at TEXTOR [12].

3.2. Short radial correlation (SRC)

The short radial correlation lengths for turbulence measured with reflectometry were analyzed in TEXTOR, it revealed some spectrum properties linked to QC modes with means of the magnitude squared coherence (MSC) [6], it has been observed that MSC allows to separate spectrum components and particularly shows a high correlation for frequencies from 50 Khz to 150 Khz and -150KHz to -50KHz corresponding to QC modes. The MSC is defined as :

$$\Gamma(f) = \frac{|G_{xy}(f)|^2}{G_{xx(f)}.G_{yy(f)}} \tag{16}$$

where $G_{xy}(f)$ denotes cross spectral density and $G_{xx}(f)$, $G_{yy}(f)$ auto spectral density.

For the rest of the paper, the word *coherence* is used to denote the MSC. As coherence spectrum is well-shaped for QC modes because of its SRC properties, a comparison can be made between the coherence from correlation reflectometry data at TEXTOR and the coherence using this work's method.

This method allows an extraction of QC modes coming from the signal received by one of correlation reflectometry antennas. Fig 12 shows the initial and extracted spectra for both antennas. Due to the symmetry and the wave emission pattern, a mirror effect can be seen on the spectrum.

After QC modes extraction for both antenna's signals processed separately, we perform MSC on it following a procedure described below:

For the orange curve , the signals from the 2 reflectometry antennas (approx. 80 000 data points) are split into 8 sets of 10 000 points then our method is applied to each of them. Thereafter, coherence is calculated pairwise between both antenna's sets and finally an average of all coherence spectrum is done. For the blue curve, Eq. (16) is performed between the whole signal of both antennas.

As shown in the Fig. 13, both coherence spectra follow each other in shape and amplitude, showing in this way a good agreement. This comparison confirms that our method preserves the signal properties in the extracted QC modes signal.

4. Conclusion and forthcoming studies

A first extraction of QC modes from reflectometry raw signals has been performed. QC modes and other component's properties are preserved after application of our algorithm as shown in the section *3.1* for the physical perspective and for the statistical one in *3.2*.

From this existing method we can study the dynamics and the propagation of QC modes using poloidal correlation reflectometers.

Moreover, an analysis of each extracted spectrum component can be carried out after the application of our method. The physical and statistical properties of each component can then be determined separately, and their interplay can be studied as well.

This method opens up the doors to automatize the extraction of modes for statistical analysis (big data), and to study the coupling with other spectrum components [2]. The method can first be used to detect whether QC mode contribution is important or negligible, consequently a machine learning algorithm can be built to discriminate spectrum with and without QC modes. It is also necessary to develop a self-determination of thresholds using Artificial Intelligence methods. In that regard, the screening technique is expected to be improved with the concept of threshold learning in which the system would learn from the whole database and thereby offer a better discrimination of QC modes for each signal. Likewise, clustering technique is expected to integrate another parameter in the vector space representation: the metric concept. It would led to a separation of coupling modes.

Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

[1] Truc A, Group T *et al.* 1986 *Nuclear fusion* **26** 1303
[2] Garbet X 2006 *Comptes Rendus Physique* **7** 573–583
[3] Diamond P H, Itoh S I and Itoh K 2010 *Modern Plasma Physics: Volume 1, Physical Kinetics of Turbulent Plasmas* (Cambridge University Press)
[4] Heuraux S and Clairet F 2009 *Instrumentation, Mesure, Métrologie* **9** 59–86
[5] Vershkov V, Shelukhin D, Soldatov S, Urazbaev A, Grashin S, Eliseev L, Melnikov A *et al.* 2005 *Nuclear fusion* **45** S203
[6] Krämer-Flecken A, Dreval V, Soldatov S *et al.* 2004 *Nuclear fusion* **44** 1143
[7] Arnichand H, Kraemer-Flecken A, Hacquin S and Sabot R 2013 Quasi coherent modes in textor and tore supra tokamaks *Proc. 11th Intl. Reflectometry Workshop-IRW11 (Palaiseau)*
[8] Arnichand H, Citrin J, Hacquin S, Sabot R *et al.* 2015 *Plasma Physics and Controlled Fusion* **58** 014037
[9] Zhong W, Shi Z, Yang Z, Xiao G *et al.* 2016 *Physics of Plasmas* **23** 060702
[10] Rice J, Citrin J, Cao N, Diamond P, Greenwald M and Grierson B 2020 *Nuclear Fusion* **60** 105001
[11] Arnichand H 2015 *Identication of trapped electron modes in frequency fluctuation spectra of fusion plasmas* Ph.D. thesis Aix Marseille Université

[12] Krämer-Flecken A, Soldatov S, Xu Y, Arnichand H, Hacquin S, Sabot R, team T *et al.* 2015 *New journal of physics* **17** 073007
[13] Sun Y, Sabot R, Hornung G, Heuraux S, Hacquin S and Verdoolaege G 2018 *Review of Scientific Instruments* **89** 073504
[14] Sun Y, Sabot R, Heuraux S, Garbet X, Hacquin S, Hornung G and Verdoolaege G 2019 *Physics of Plasmas* **26** 032307
[15] Mallat S 1999 *A wavelet tour of signal processing* (Elsevier)
[16] Schubert E, Sander J, Ester M, Kriegel H P and Xu X 2017 *ACM Transactions on Database Systems (TODS)* **42** 1–21
[17] Vasiliev V V and Fedorov L V 2018 *Journal of Modern Physics* **9** 2482–2494
[18] Kaiser G 2011 Continuous wavelet transforms *A Friendly Guide to Wavelets* (Springer) pp 60–77
[19] Daubechies I 1990 *IEEE transactions on information theory* **36** 961–1005
[20] Farge M 1992 *Annual review of fluid mechanics* **24** 395–458
[21] Torrence C and Compo G P 1998 *Bulletin of the American Meteorological society* **79** 61–78
[22] Li Y and Wu H 2012 *Physics Procedia* **25** 1104–1109
[23] Sculley D 2010 Web-scale k-means clustering *Proceedings of the 19th international conference on World wide web* pp 1177–1178
[24] Arthur D and Vassilvitskii S 2006 k-means++: The advantages of careful seeding Tech. rep. Stanford
[25] Sabot R, Sirinelli A, Chareau J M and Giaccalone J C 2006 *Nuclear fusion* **46** S685
[26] Sung C, White A, Howard N, Oi C *et al.* 2013 *Nuclear Fusion* **53** 083010
[27] Grierson B, Chrystal C, Haskey S, Wang W *et al.* 2019 *Physics of Plasmas* **26** 042304
[28] Angioni C, Peeters A, Ryter F, Jenko F, Conway G, Dannert T, Fahrbach H, Reich M, Suttrop W, Team A U *et al.* 2005 *Physics of plasmas* **12** 040701
[29] Conway G, Angioni C, Dux R, Ryter F *et al.* 2006 *Nuclear fusion* **46** S799
[30] Rice J, Gao C, Reinke M, Diamond P *et al.* 2013 *Nuclear Fusion* **53** 033004
[31] Erofeev I, Fable E, Angioni C, McDermott R *et al.* 2017 *Nuclear Fusion* **57** 126067
[32] Lee W, Lee J, Lee D J, Park H K *et al.* 2020 *Nuclear Fusion* **61** 016008
[33] Arnichand H, Sabot R, Hacquin S, Krämer-Flecken *et al.* 2014 *Nuclear Fusion* **54** 123017

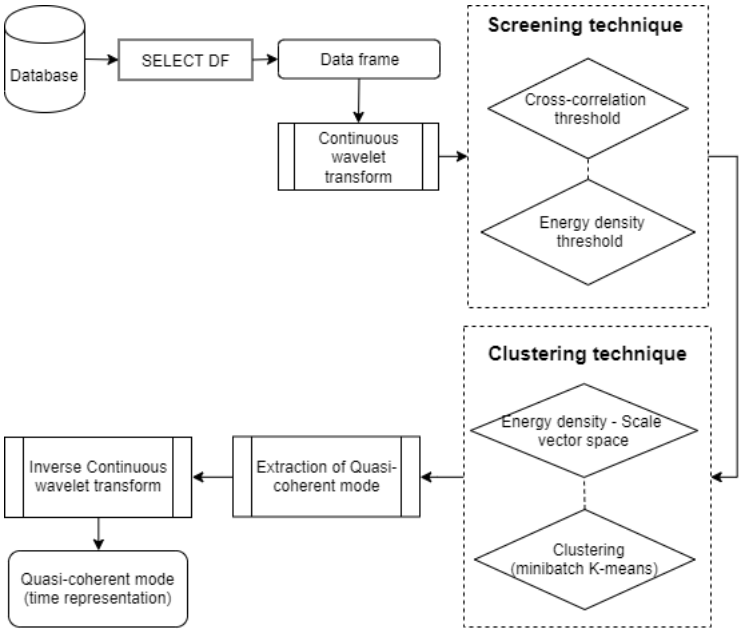

Figure 1. Data-flow diagram: rounded rectangle for Data, rectangle double line for Process and rhombus for proposed techniques

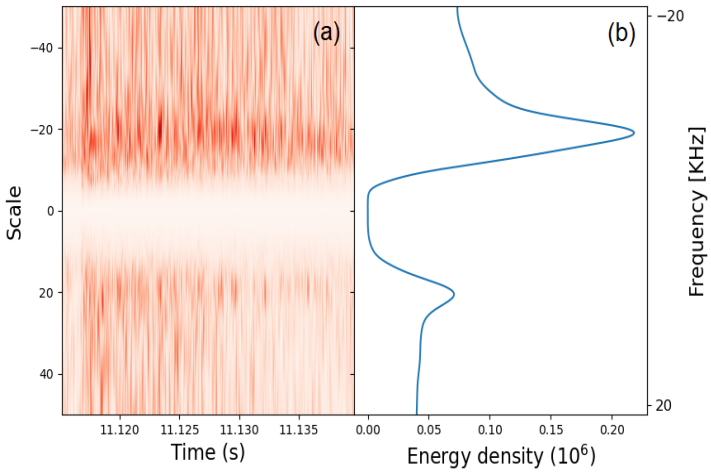

Figure 2. CWT modulus of $f(t)$ (a) and the averaged energy density as a function of scales (b)

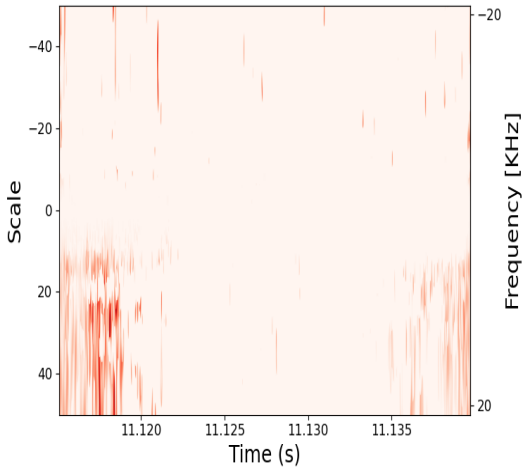

Figure 3. $|Wf(u,s)|$ after cross-correlation threshold (T_{h1})

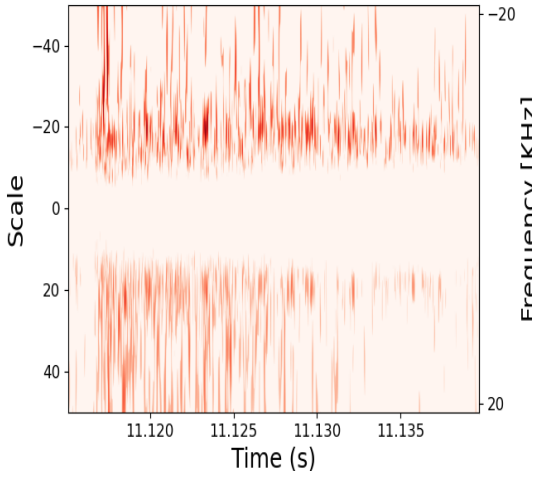

Figure 4. $|Wf(u,s)|$ after energy density threshold (T_{h2})

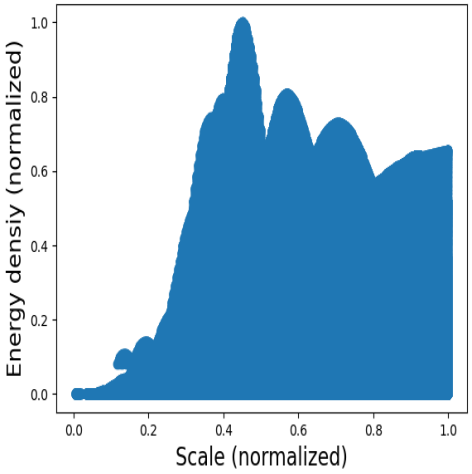

Figure 5. Projection of $Wf(u,s)_{screened}$ into energy density-scale vector space

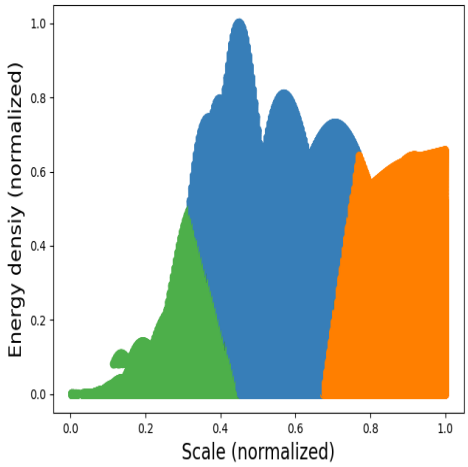

Figure 6. Clustering technique on **V** by Minibatch K-means

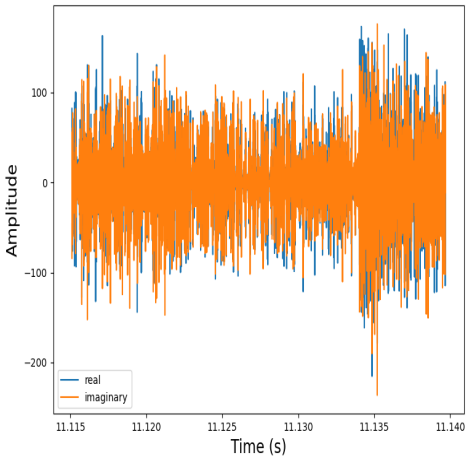

Figure 7. Time representation of Quasi-coherent modes

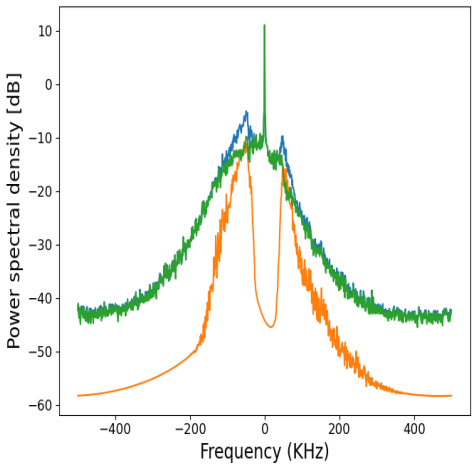

Figure 8. Frequency spectrum of initial signal (blue curve), QC modes (orange curve) and signal after extraction of QC modes (green curve)

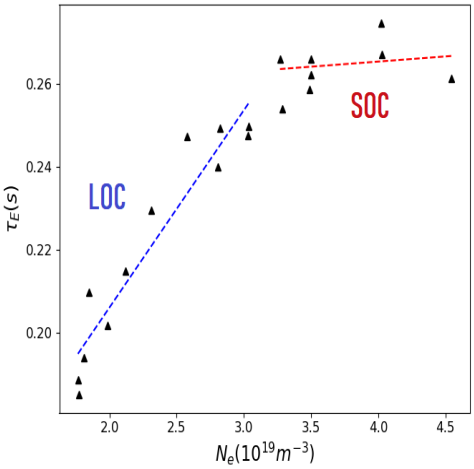

Figure 9. Tore Supra confinement time versus density during the LOC-SOC transition at $I_p = 1.2MA$ for $r/a = 0.3 \pm 0.025$ and $B_t = 3.8T$ from the shot range of #41003 to #41013

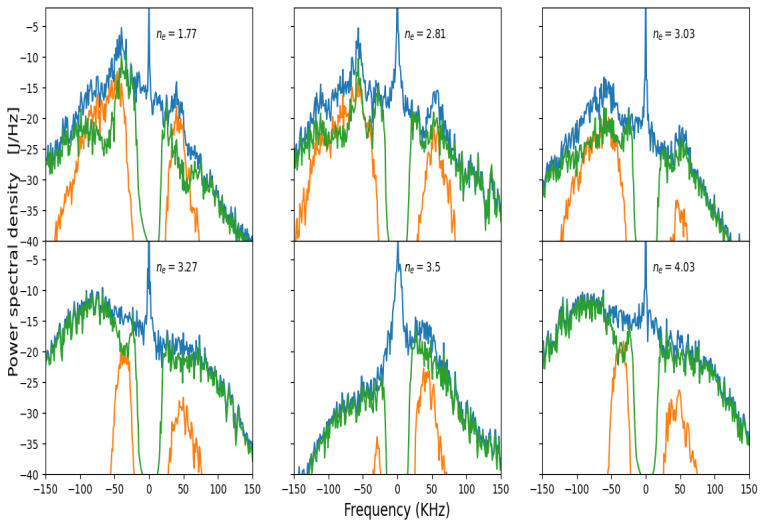

Figure 10. Tore Supra reflectometry spectra during LOC-SOC transition at $I_p = 1.2MA$ for $r/a = 0.3 \pm 0.025$ and $B_t = 3.8T$ from the shot range of #41003 to #41013

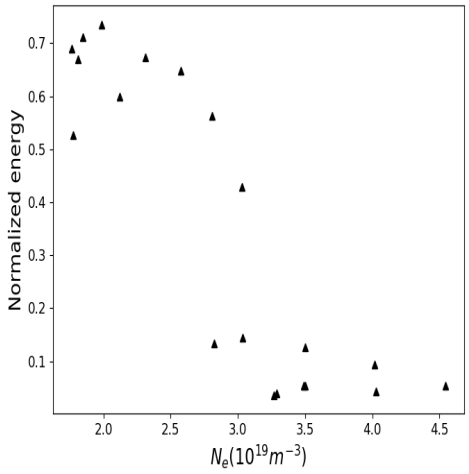

Figure 11. Normalized energy of QC modes as a function of electron density

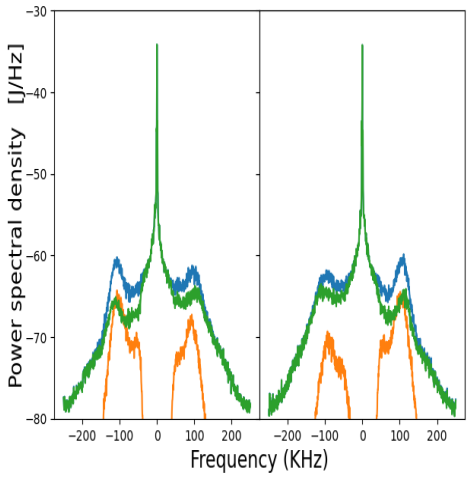

Figure 12. Frequency spectrum of initial signal from both antennas (blue curve), QC modes (orange curve) and signal after extraction of QC modes (green curve)

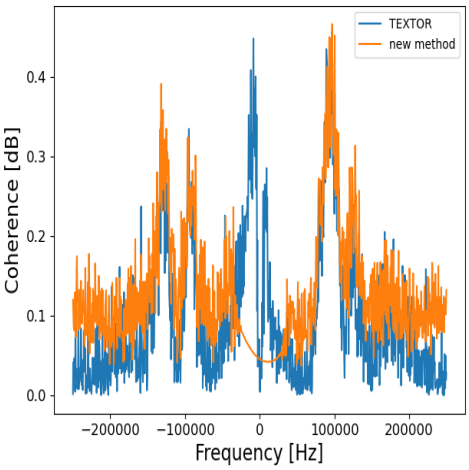

Figure 13. Coherence spectrum: TEXTOR method (blue curve) and proposed method (orange curve)