OMFIT-Fluctana Tutorial

DISCLAIMER:

This is an editable public document! When you make a change to it, it will be visible immediately to everyone! Feel free to edit it and help other OMFIT users!

Since this is an evolving document, there may be some small inconsistencies as different figures have been taken by different people with different versions of OMFIT for different analyses.

Minjun J. Choi and Myungwon Lee
NFRI, Korea

Introduction

The OMFIT-Fluctana module is developed to provide spectral and statistical analysis methods for fluctuation data from tokamak experiments and simulations.

GUI design of the OMFIT-Fluctana module : Dr. Minjun J. Choi (NFRI, Korea) and Mr. Myungwon Lee (NFRI, Korea)
GUI programming of the OMFIT-Fluctana module : Mr. Myungwon Lee (NFRI, Korea) and Dr. Minjun J. Choi (NFRI, Korea)
Source code (‘fluctana’ Python package) programming : Dr. Minjun J. Choi (NFRI, Korea)

Release Note

The first release of the OMFIT-Fluctana version 1.0 is on June X, 2020.
The version 1.0 includes

- Spectrogram and cross spectrogram plot (short-time FFT base)
- Correlation & correlation coefficient analysis (FFT base)
- Correlation image plot
- Linear spectral analyses such as (either FFT or Morlet wavelet base)
  - Cross power spectrum
  - Cross coherence spectrum
  - Cross phase spectrum
  - Local wavenumber and frequency spectrum
- Nonlinear spectral analyses such as (either FFT or Morlet wavelet base)
  - Bicoherence
- Statistical analyses such as
  - Skewness and Kurtosis
  - The Hurst exponent
  - The Jensen-Shannon Complexity and the normalized Shannon Entropy
  - Intermittency

Future updates will include analyses of nonlinear spectral energy transfer and entropy transfer. Although the OMFIT-Fluctana module is designed as a universal tool for data from various tokamak devices and simulations, in version 1.0 users are only able to do analyses using the KSTAR basic diagnostics data (saved in the KSTAR MDSplus server) and the KSTAR ECEI data on the iKSTAR or uKSTAR server. More diagnostics or simulation data will be accessible in a later version.

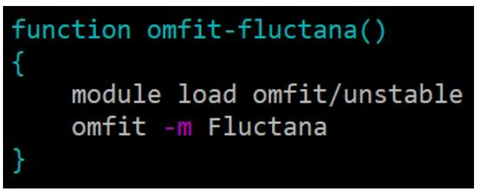
Getting Started

The OMFIT-Fluctana module is based on the 'fluctana' Python package (https://github.com/minjunJchoi/fluctana, developed by Dr. Minjun J. Choi (NFRI, Korea)). Users have to add a 'fluctana' path to their Python path before using the OMFIT-Fluctana module. For example, users on the uKSTAR server can copy and paste the following line into their `~/.bashrc` file. `/code/fluctana` is the 'fluctana' git clone on the uKSTAR server.

```
export PYTHONPATH=/code/fluctana:$PYTHONPATH
```

It would be helpful if users could make a function in `~/.bashrc` as follows to call the OMFIT-Fluctana module directly at the terminal.

```
function omfit-fluctana ()
{
      module load omfit/unstable
      omfit -m Fluctana
}
```



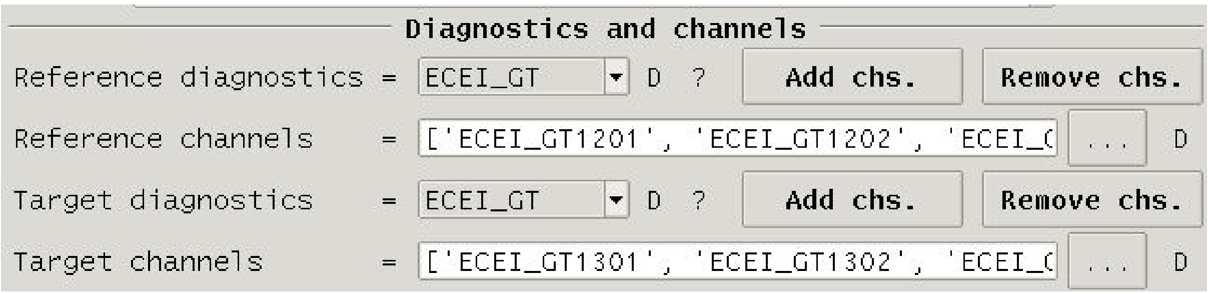
Basic Structure of GUI

The OMFIT-Fluctana GUI is developed to provide an easy way to run various analyses for many channels of data. The GUI can be divided into 4 panes. In the first three panes (from the top), users would set diagnostics parameters, basic parameters, and (optional) filtering parameters. In the last pane, users select an analysis tab to run and set analysis parameters if necessary and click the 'PLOT' buttons to see the result.

1. Setting diagnostics parameters

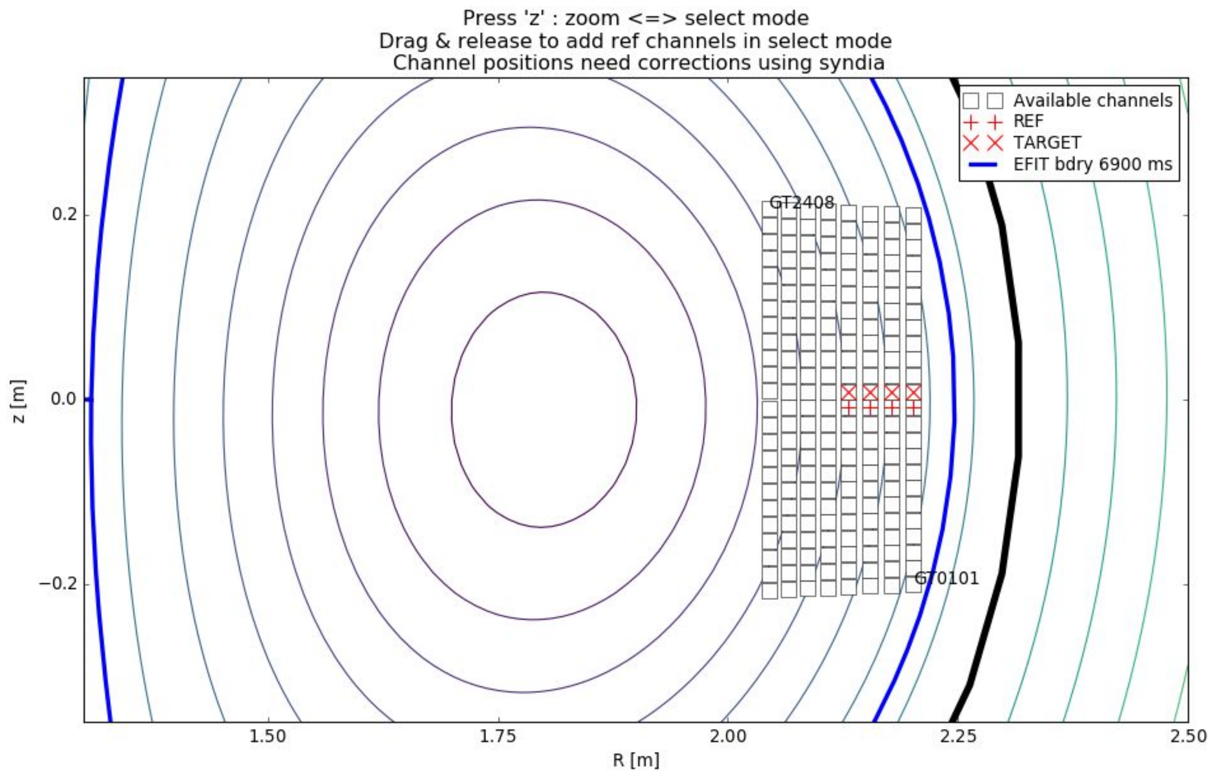
Users can first select ‘Device’ and ‘Shot’ as shown in the following figure.

Next, choose ‘Reference/Target diagnostics’ and type in a list of the channel name. Users can use the MDSplus node name to access some data via the MDSplus server, or users need to know the proper channel name for a particular diagnostic.



For some multi channel diagnostics such as the KSTAR ECEI diagnostic, users can click the ‘Add chs.’ or ‘Remove chs.’ button to add or remove channels easily from the list shown in the ‘Reference/Target channels’ entry.

The figure shown below will appear if users click the button and users can drag and release to add or remove channels. If users press ‘z’, users can zoom in.



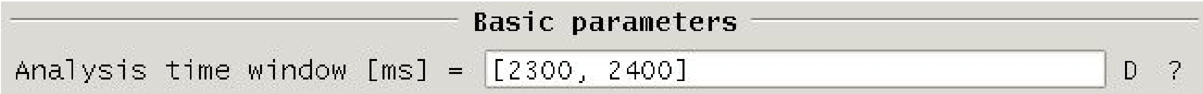
Or, users can use the contracted form such as ‘ECEI_GT1301-1304’ to add ‘1301’, ‘1302’, ‘1303’, and ‘1304’ if the diagnostic module supports this.

Please note that a careful selection of channels is important to get a meaningful result. Correlation and spectral analyses require both ‘Reference channels’ and ‘Target channels’, while statistical analyses require only ‘Target channels’. Please refer to the [Learning with Examples]() section for detailed instruction for each type of analysis.

2. Setting basic parameters

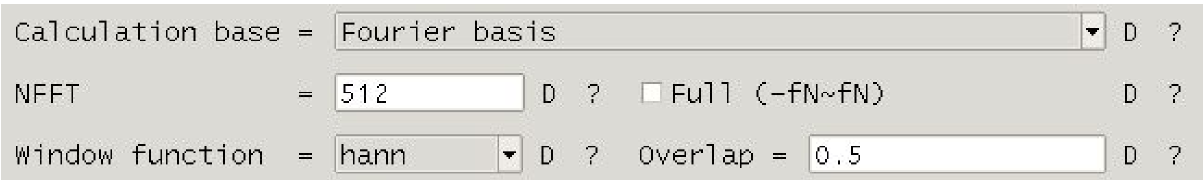
Basic parameters include ‘Analysis time window’, ‘Detrend option’ and ‘Normalization option’, and parameters for the Fast Fourier Transform (FFT) or the wavelet transform for correlation and spectral analyses.

‘Analysis time window’ means a period over which the data is read and the analysis will be done using that data.

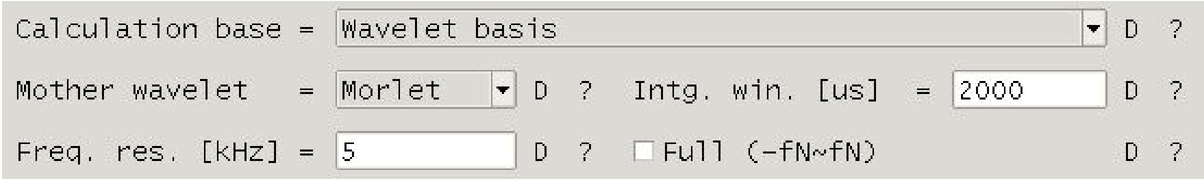


‘Detrend option’ has 3 choices: ‘none’ = do nothing, ‘mean’ = subtract a time average from the data, ‘linear fit’ = subtract a linear fit from the data.
‘Normalization option’ has 3 choices: ‘none’ = do nothing, ‘mean’ = normalize the data by its time average, ‘low-pass’ = normalize the data by its low-frequency component.

If users select ‘Fourier basis’ for the calculation base of correlation or spectral analyses, users will see the figure below and set up a few more inputs. ‘NFFT’ determines the frequency resolution of spectral analyses, and it is related to the accuracy of the calculation (the number of ensembles for the calculation). A higher ‘NFFT’ value results in a finer frequency resolution but with more uncertainty (with a lower number of ensembles). Check ‘Full’ if users are working with the complex valued diagnostics data or doing some analyses which require the full spectrum. Since the ‘fluctana’ spectral analyses follow Welch’s approach [1], users have to indicate which ‘Window function’ will be applied in the FFT. ‘hann’ (Hanning window) with ‘Overlap’ ratio 0.5 will be fine for most cases, but users can try different windows.

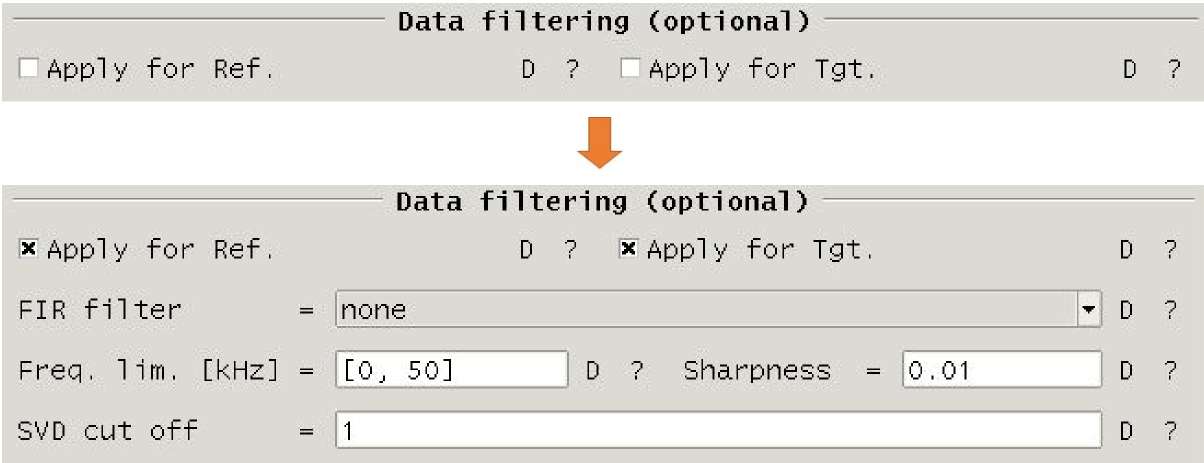


If users select ‘Wavelet basis’ for the calculation base of spectral analyses, users will see the figure below and set up a few more inputs. The ‘Wavelet basis’ option is useful when users are required to do spectral analyses for a short time window [2]. ‘Morlet’ for ‘Mother wavelet’ will be fine for most cases. ‘Integration window’ takes a period (in microsecond!) over which a time average of wavelet transform W(t, f) is done to provide W(f). Please note that ‘Analysis time window’ of a few times of ‘Integration window’ is enough for the analysis. Since the wavelet base calculation of ‘fluctana’ is based on the continuous wavelet transform, users can specify the ‘Frequency resolution’ freely, but care must be taken. The wavelet base calculation can provide spectral analyses with a better temporal resolution than the FFT base calculation, but its spectral frequency information leaks more.



3. Setting filtering parameters (optional)

Users can apply filters on data before doing any analysis. The GUI for setting filtering parameters will appear if users check ‘Apply for Reference/Target’.



Finite impulse response filters (‘FIR filter’) will be used to get low pass, band pass, high pass, or band block filtered data from the raw signal. For example, ‘FIR filter’ = ‘FIR pass’ and ‘Frequency limit’ = [0, 50] means a low pass filtering < 50 kHz, and ‘FIR filter’ = ‘FIR pass’ and ‘Frequency limit’ = [50, 0] means a high pass filtering > 50 kHz. ‘Sharpness’ determines the sharpness of the frequency cut-off of the filter.

If users want to apply additional filters based on singular value decomposition (SVD), users can type in a ‘SVD cut off’ value other than 1. ‘SVD cut off’ means the power fraction of the remaining components in the filtered data against the total power or the raw data.

4. Running analysis

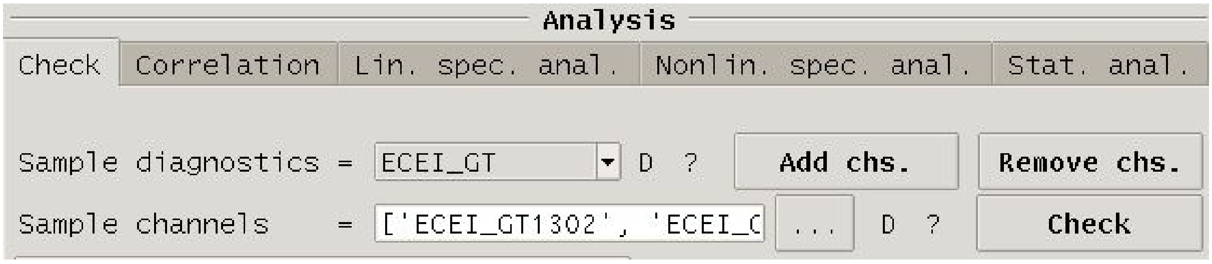
Users are now ready to run analysis. There are 5 tabs in the analysis pane. The first ‘Check’ tab is for investigating the raw data and its spectrogram to determine the ‘Analysis time window’ of the basic parameters. The ‘Correlation’ tab provides plots of the cross correlation and correlation coefficient between the reference and target channels and images of the cross correlation coefficient of the target channels against a single reference channel. The ‘Linear spectral analyses’ tab provides plots of cross power, cross coherence, and cross phase between the reference and target channels. Those spectra can be also used to obtain images of cross RMS, summed coherence, and average phase velocity within the ‘Frequency range’. The local wavenumber-frequency spectrum can be obtained using pairs of the reference and target channels. The ‘Nonlinear spectral analyses’ tab provides bicoherence analysis. The ‘Statistical analyses’ tab includes plots of Skewness and Kurtosis, Hurst exponent, Complexity and Entropy, and intermittency. These statistical methods only use the target channels.

Users can get plots of the analysis result by clicking the ‘PLOT’ buttons. If users click the 'PLOT' button without closing the previous figure, the next plots will be overlaid on the existing figure. Users may need to set some analysis parameters before clicking the ‘PLOT’ buttons. Please refer to the next section for how to run each analysis method properly.

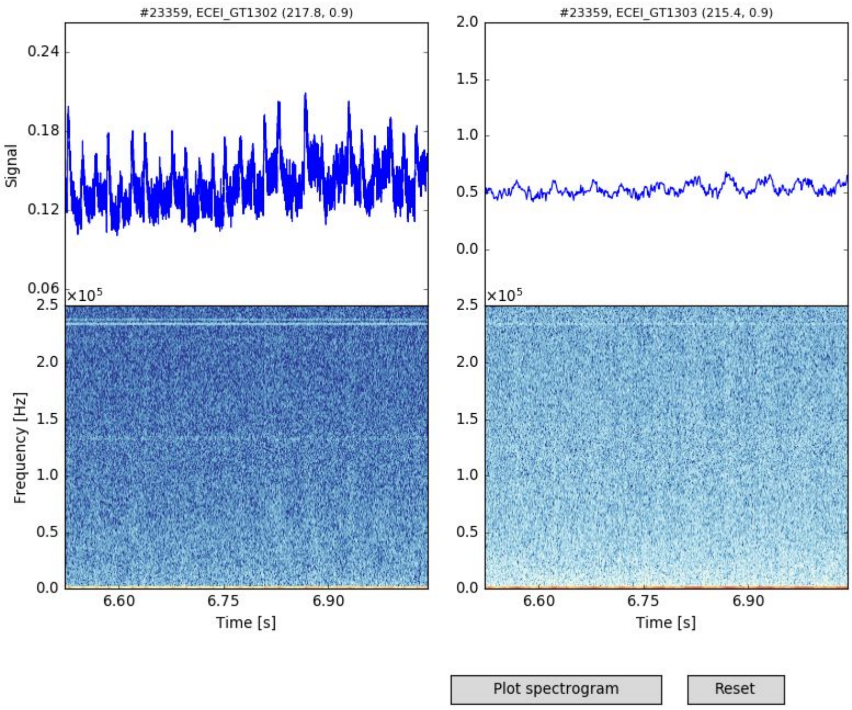
Learning with Examples

1. Check spectrogram and cross power spectrogram

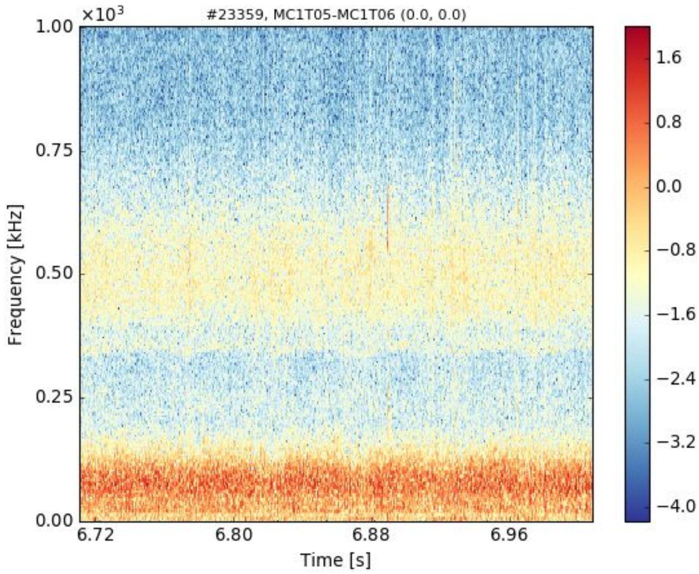
In the ‘Check’ tab, users can check the raw data of the ‘Sample channels’ of the ‘Sample diagnostics’ by clicking the ‘Check’ button. ‘Sample channels’ take at most 3 channels.



Users would see the entire (down-sampled) signal in the top axes. Users can zoom-in the view and check the spectrogram by clicking the ‘Plot spectrogram’ button.



The result shows that no clear MHD activities (which usually appear as well defined lines) are found in the spectrogram of the ECEI data (electron temperature fluctuations). Note that the end points of the current axes will be taken as the ‘Analysis time window’ of the ‘Basic parameters' when the figure is closed. Users may also want to check the electromagnetic fluctuations via the cross power spectrogram between the KSTAR Mirnov coils data 'MC1T05' and 'MC1T06' over the 'Analysis time window' by clicking the 'Cross power spectrogram between' button.

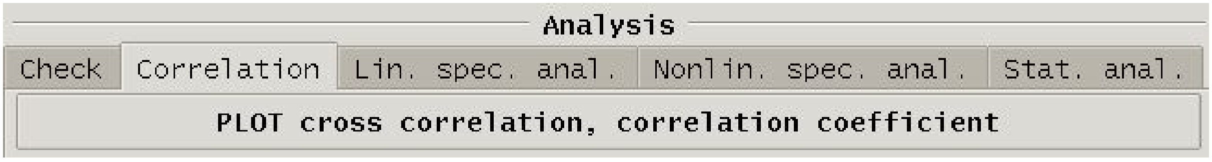


Found no MHD activities, but the broadband fluctuations may need to be more investigated as follows.

2. Correlation analyses

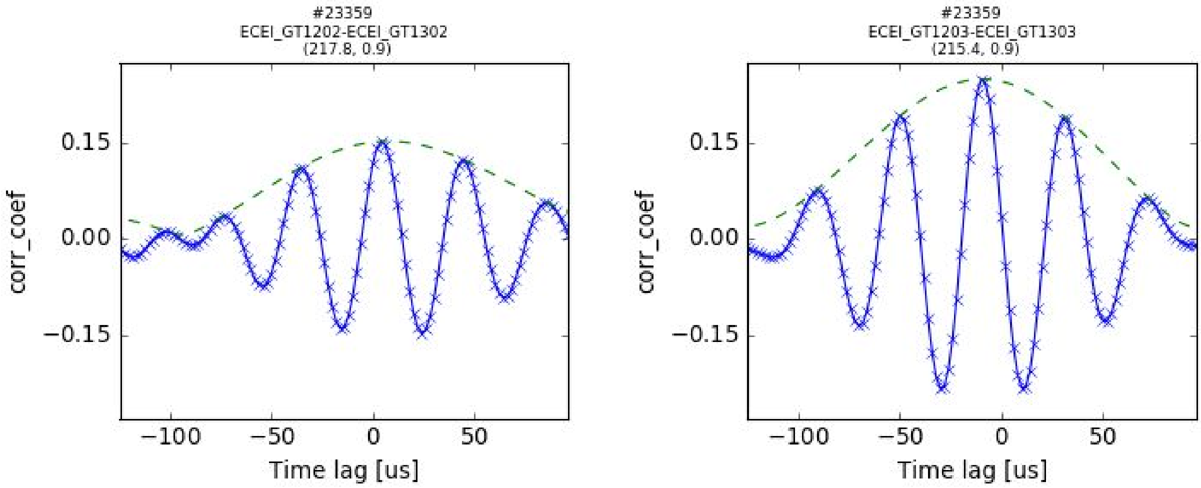
INSTRUCTION for selecting channels: users should have at least one reference channel to run this analysis, since it is calculating the correlation function between the reference data and target data. If users add the same number of reference and target channels, the calculation will be done for each pair of reference and target channels made by an order in the channel list. For example, in the case of 'Reference channels' = ['ECEI_GT1202', 'ECEI_GT1203'] and 'Target channels' = ['ECEI_GT1302', 'ECEI_GT1303'], one calculation will be done using 'ECEI_GT1202' and 'ECEI_GT1302' and the other using 'ECEI_GT1203' and 'ECEI_GT1303'. For the ECEI diagnostics, '1202' and '1302' means a pair of vertically adjacent channels at the radial position '02', and so the two resulting plots will show the correlation function between vertically adjacent channels (or, vertically coherent signals) at two radial positions '02' and '03', respectively. On the other hand, if users add a single channel for 'Reference channels', the calculations between one reference data and all other target data will be done.

Users can check cross correlation and correlation coefficient (normalized correlation) by just clicking the following 'PLOT' button.



No analysis parameters are required, but since the cross correlation is often used to measure the correlation time or correlation length of a particular fluctuation component, it is desired to apply the band pass filter before running the analysis.

The plots below are the correlation coefficient (2nd tab) obtained with the band pass (20--30 kHz) filtered data.



They indicate that

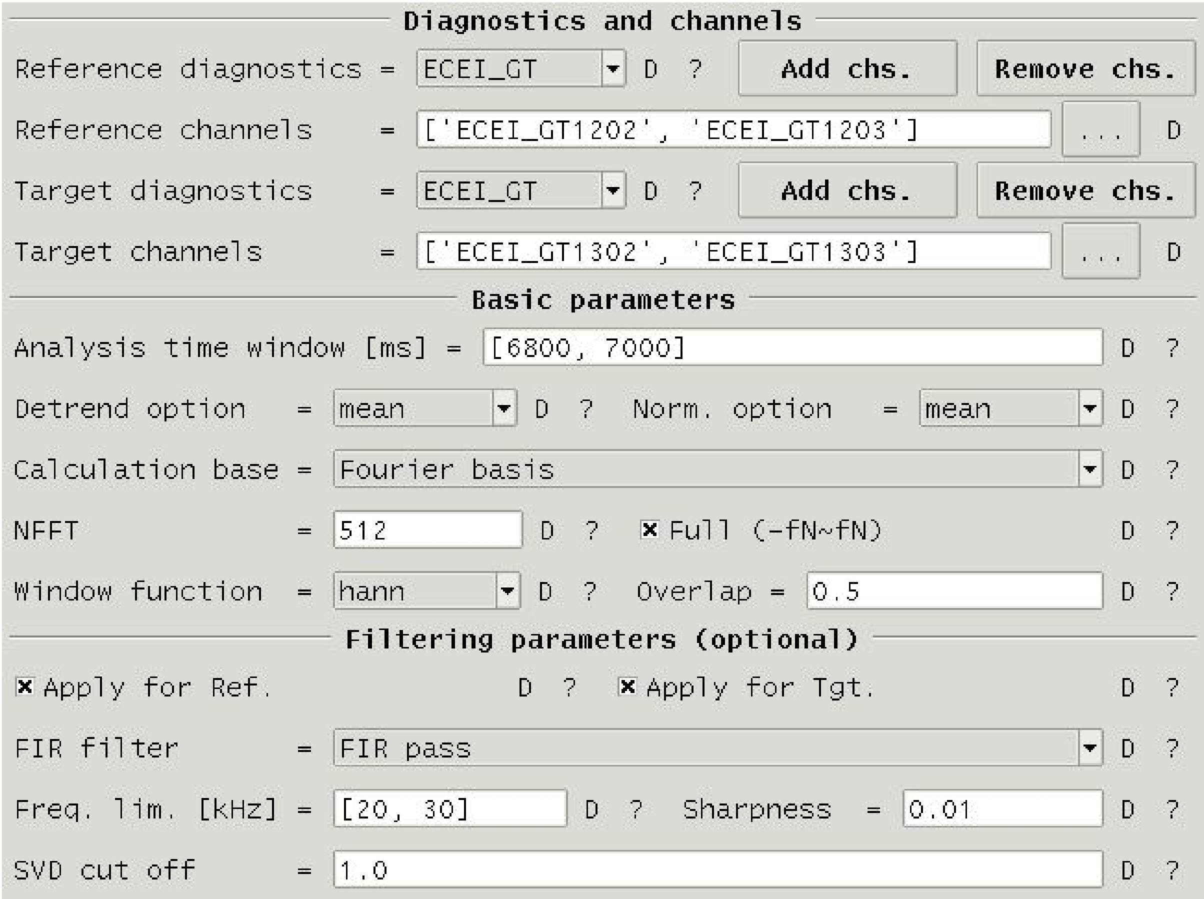
1) the fluctuation is more coherent vertically at ‘03’ than ‘02’.

2) the structure making this fluctuation is moving from ‘12’ to ‘13’ (upwards; the positive phase velocity) at the radial position ‘02’ (positive time lag at the maximum correlation coefficient = ‘1302’ is the delayed signal) while it is moving from ‘13’ to ‘12’ (downwards; the negative phase velocity) at the radial position ‘03’ (negative time lag at the maximum correlation coefficient = ‘1303’ is the preceding signal).

3) the fluctuation has a cross correlation time of ~100 us.

The first two results would be more clearly seen in linear spectral analyses as shown in the next section.

Please refer to the following figure for parameters used to make the above result.

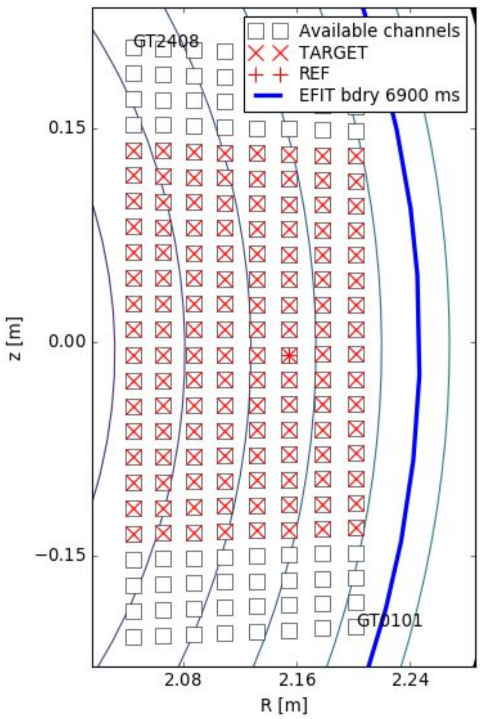


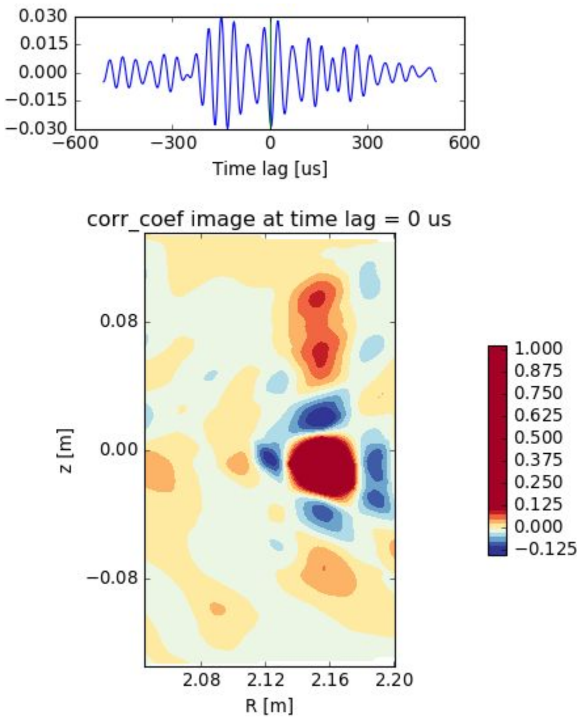
Note that ‘Full’ should be checked for the correlation calculation (it is hard-coded anyway).

Users can estimate the correlation length of the fluctuation by clicking the following ‘PLOT’ button. ‘Sample channel number’ means a channel index in the target channel list whose correlation coefficient will be shown at the top axes.

The calculation of the correlation coefficient image needs only one reference channel and many target channels to be compared as shown below.

The result figure shows an image of cross correlation coefficients at the given time lag (indicated by the green line in the top axes) between the reference channel data and all target channels data. Users can change the time lag by clicking at the top axes and the figure will be closed if the users' click is out of the range.



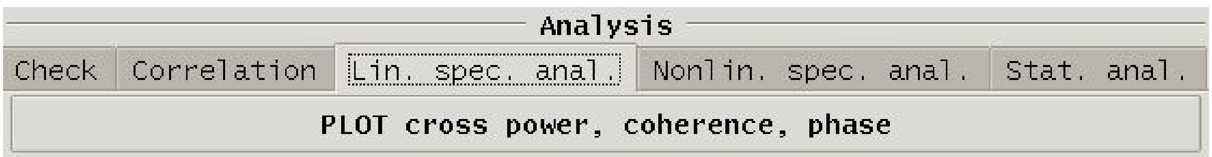


The correlation length can be estimated on the image with the zero time lag. It shows that the radial and poloidal correlation length of the fluctuation (20--30 kHz) at the reference channel position is about ~5 and ~10 cm (similar to the wavelength in each direction), respectively. Note that this finite correlation length affects the estimation of the correlation time of the moving structure.

3. Linear spectral analyses

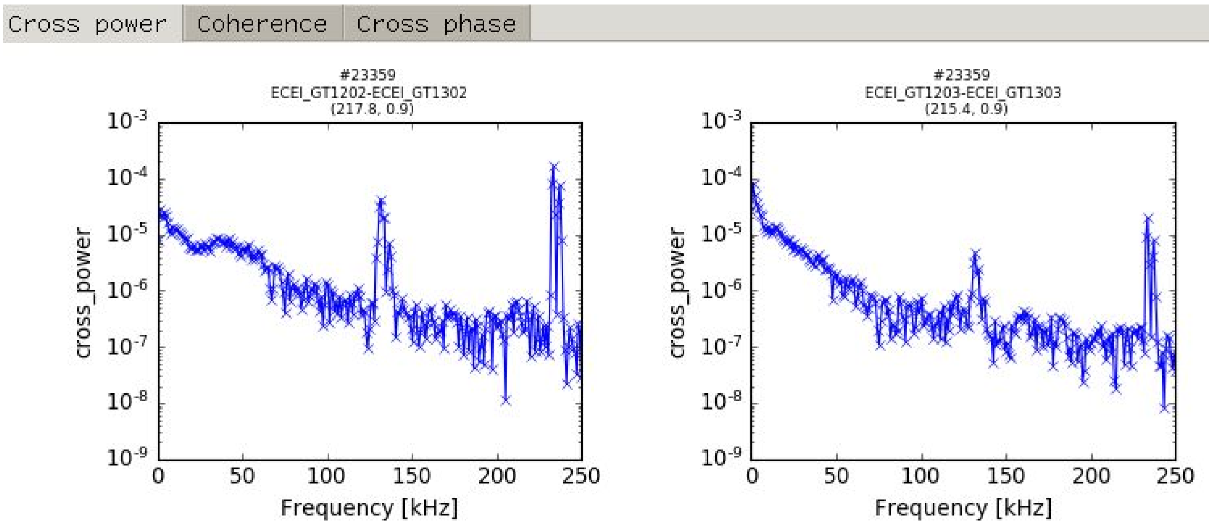
INSTRUCTION for selecting channels: it would be the same as that of the correlation analyses except that the local wavenumber and frequency spectrum requires a set of adjacent pairs along the desired wavenumber direction. More explanation on this will be given later.

Users can check the cross power, cross coherence (normalized cross power), and cross phase between pairs of reference and target channels by clicking the following 'PLOT' button.

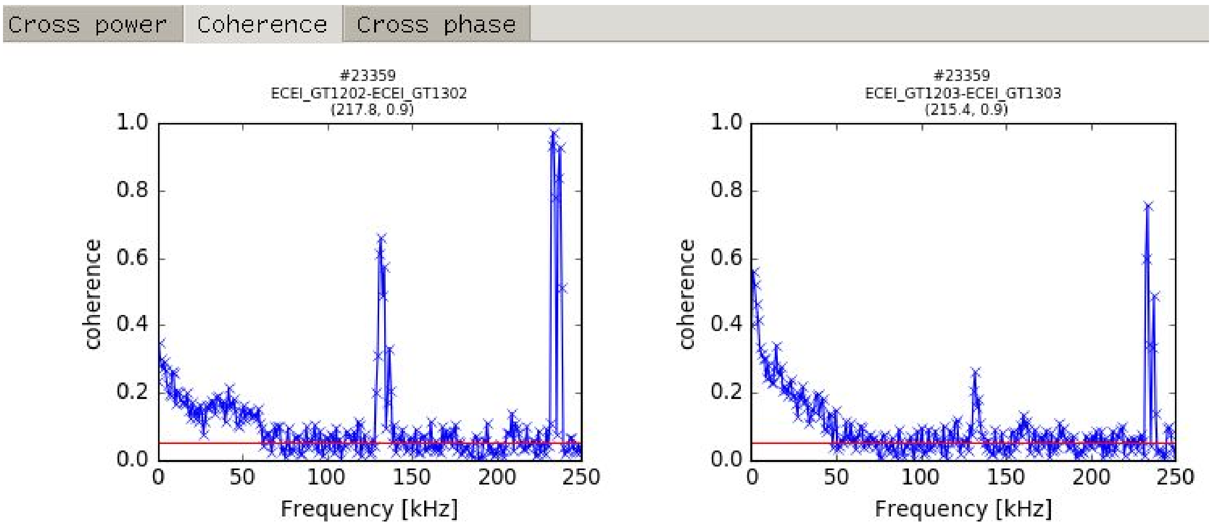


No data filtering is required for this analysis, but the data normalization is often essential since the many-channel diagnostics are rarely absolutely calibrated.

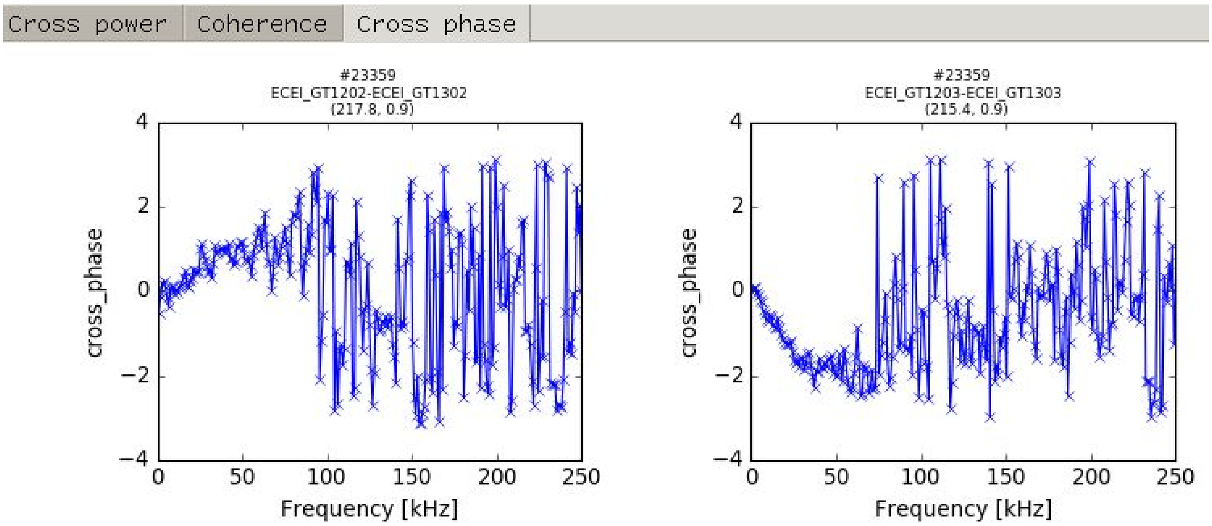
The resulting figure has three tabs. The cross power is useful to measure the quantitative amplitude of the fluctuation, but it includes the noise background. The cross coherence shows the normalized cross power with a significance level, and so it is useful to check the existence of coherent fluctuation beyond the (incoherent) noise level. The cross phase means a phase difference between reference and target channels. It represents the local dispersion relation for spatially separated channels measuring the same quantity. But, if users use channels measuring different physical quantities at the same location, it is just a phase difference between the different quantities. In the following examples, the KSTAR ECEI channels measuring the electron temperature at separated locations are used (same as those used for the cross correlation plots (not the image plot)).



The cross power plots show that the fluctuation power in 0--30 kHz range is stronger at the radial location '03' than '02', but the fluctuation power in 30--60 kHz is stronger at ‘02’ than ‘03’. Strong peaks near 130 kHz and 240 kHz are the external coherent noise.



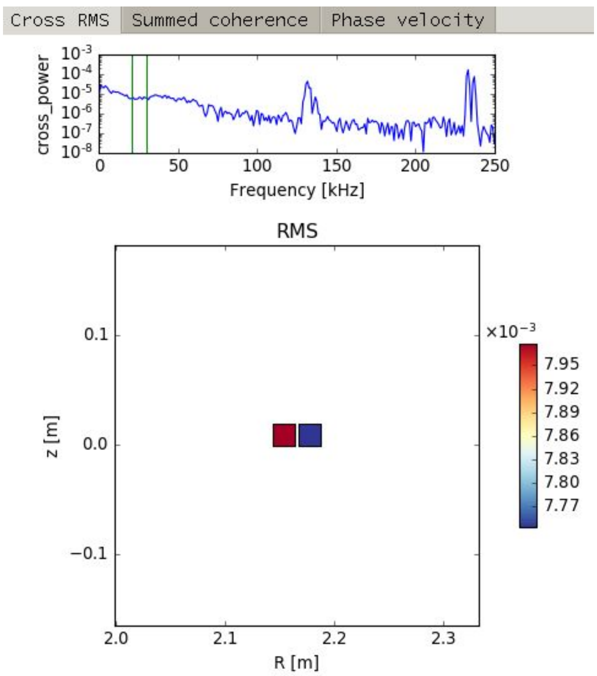
The cross coherence plots show that the coherent fluctuation exists in 0-60 kHz range at both ‘02’ and ‘03’ locations.

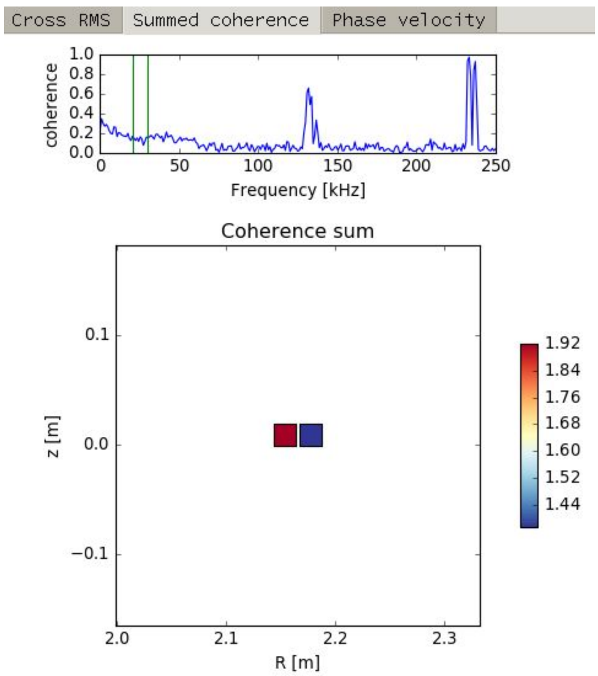


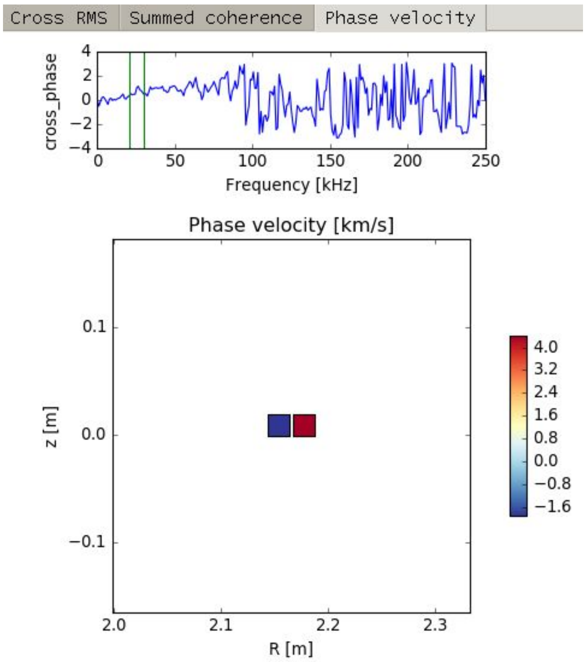
The cross phase plots show that the phase velocity is positive at '02' and negative at '03'. Note that the phase velocity = ω/k is proportional to the inverse of the slope of the above plots.

Users can estimate cross RMS, summed coherence, and phase velocity within the given 'Frequency range' based on the previous spectra by clicking the following 'PLOT' button.

Below are the results for ‘Frequency range’ = [20, 30].

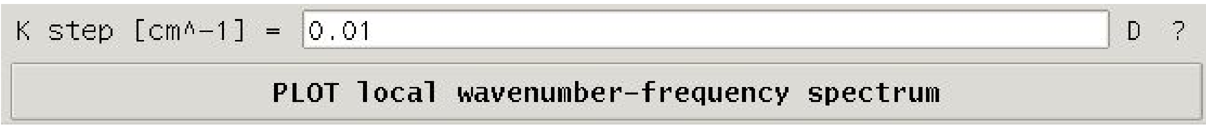




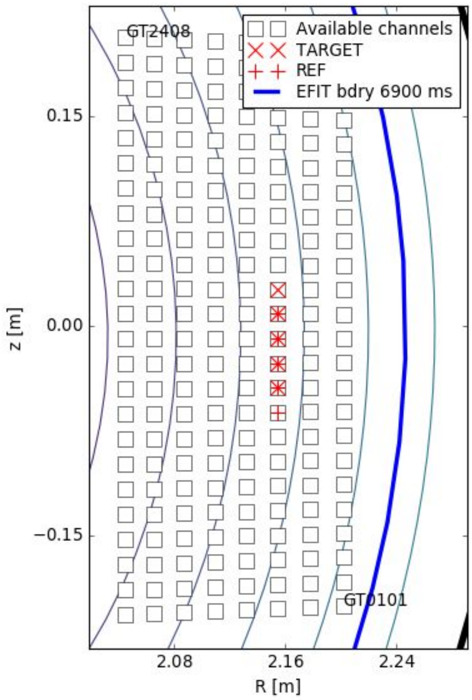


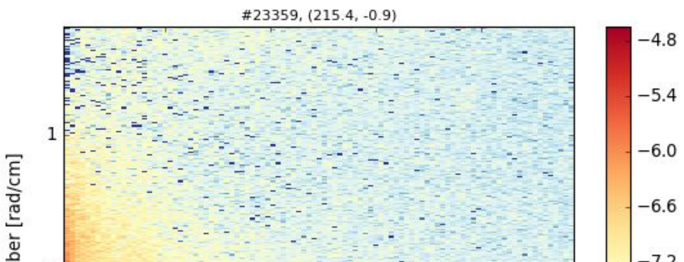
The color scale is set automatically using standard deviation of the data.

Users can check the local wavenumber and frequency spectrum by clicking the following 'PLOT' button.



This method is based on Beall's method [3]. The local wavenumber of each frequency component can be estimated by the cross phase between a pair of spatially separated channels along the desired direction. Cross phase measurements from many ensembles can be used to make a histogram to see the distribution of fluctuation power in the wavenumber and frequency space. 'K step' indicates the local wavenumber resolution of the histogram. Note that the previous cross phase plot is a result of the ensemble average of those measurements. If users want to measure the 'poloidal' local wavenumber and frequency spectrum, every pair of reference and target channels should be a pair of the 'poloidally' adjacent channels. The right figure shows one example. That intends to measure the poloidal local wavenumber and frequency spectrum at the radial position '03' (the third radial position from the right).



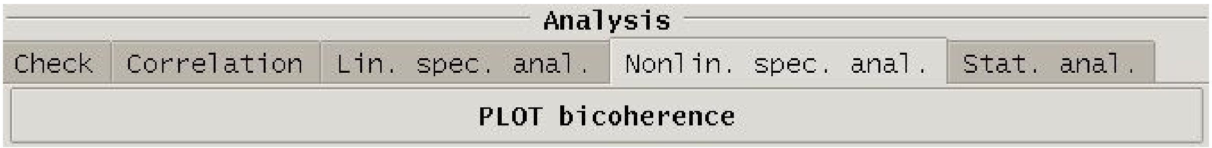


The local wavenumber and frequency spectrum can be more informative than the cross phase plot since it provides a spectral width. But, it is not easy to distinguish the noise contribution from the real spectral width, and also (as users can see from the channel selection figure) the channels are not in line with the desired poloidal direction exactly. Using multiple pairs allows an averaging to reduce the noise contribution though.

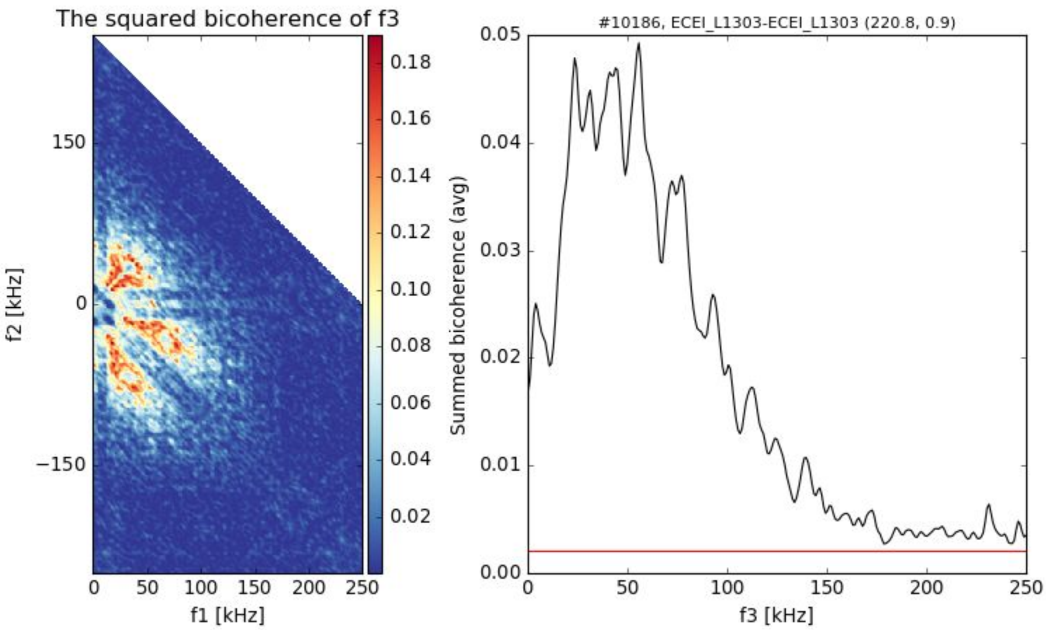
4. Nonlinear spectral analyses

INSTRUCTION for selecting channels: Since this method is investigating a higher level of information, taking more time, it is suggested that users add a fewer pairs of reference and target channels. Bicoherence [4] may require a pair of channels at the same location, since the (lab frame) frequency matching condition is enforced. Setting 'Reference channels' = 'Target channels' calculates auto bicoherence, and it would be the fair selection to investigate the three-wave coupling for most cases. But, if users are able to use channels measuring different quantities at the same location, cross bicoherence can provide additional information.

Users can check bicoherence by clicking the following 'PLOT' button.

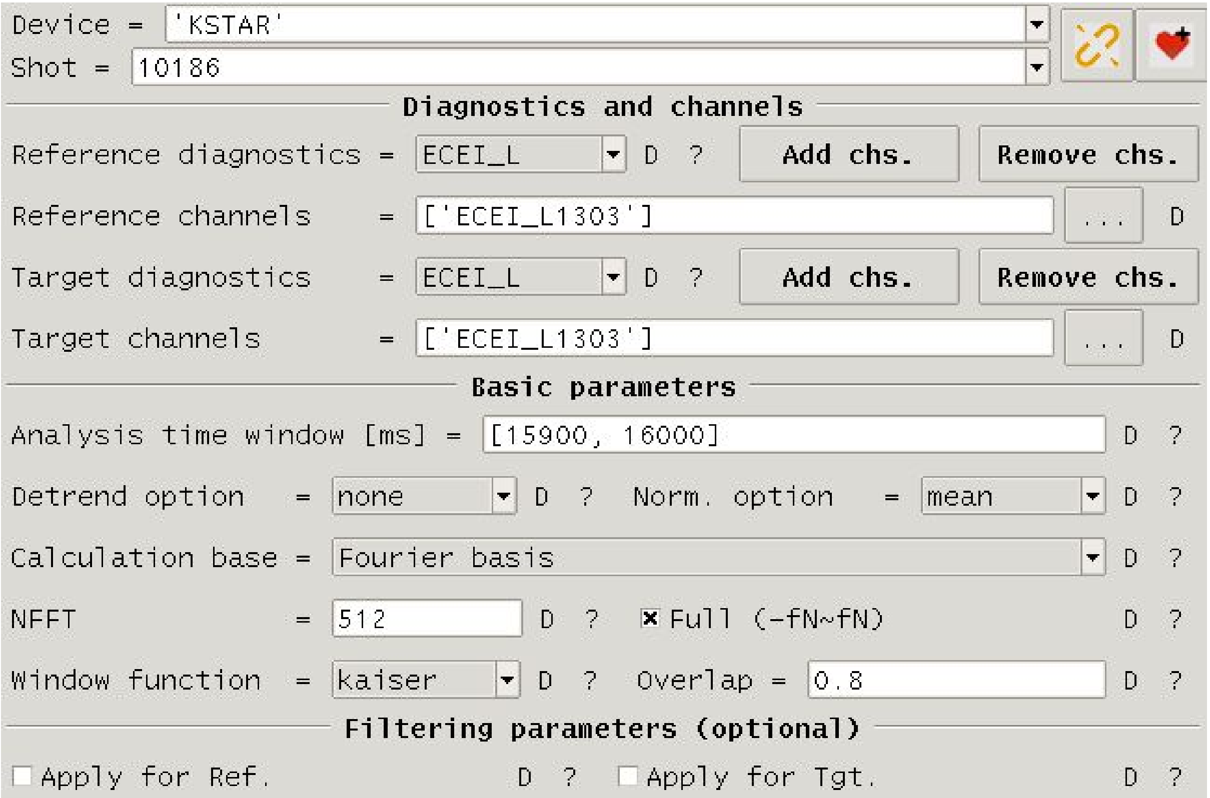


The resulting figure has two plots. The left plot shows the squared bicoherence for all combinations of f3=f1+f2. The squared bicoherence is the ratio of the fluctuation power at f3 due to the three-wave coupling with f1 and f2 and the total power at f3. The three-wave coupling is detected by the coherent phase relationship among three waves. The right plot shows the averaged bicoherence over all pairs of f1+f2 for each f3.



These show that some (2--5 %) of fluctuations in 0--100 kHz range have coherent phase relationships with two other fluctuations in the 0-100 kHz range.

The above result is obtained with



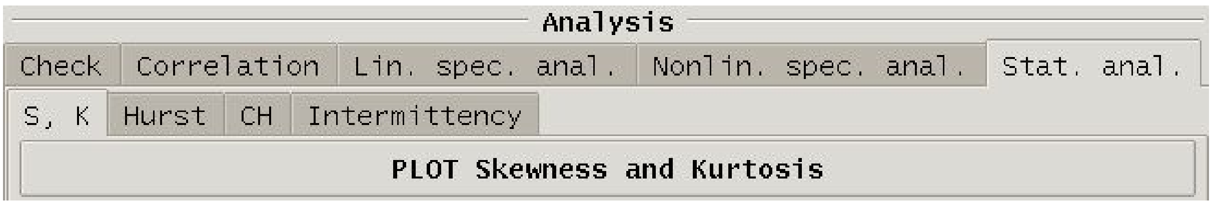
and this discharge will be used also in the following statistical analyses.

5. Statistical analyses

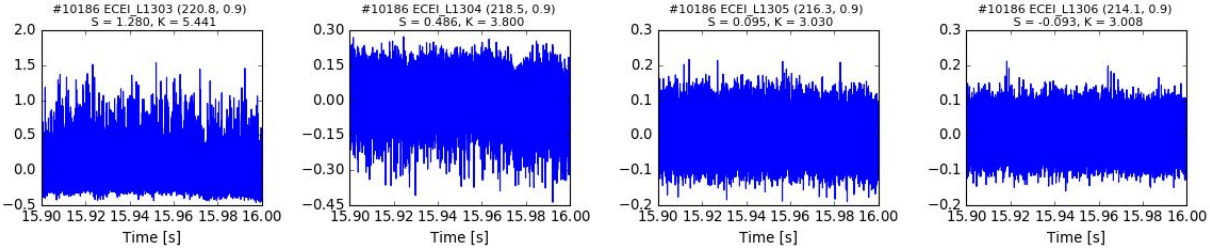
INSTRUCTION for selecting channels: These methods do not require pairs of channels. They are a single channel calculation to measure statistical quantities for each signal. Channels in the target channel list will be used.

There are four sub-tabs for four different methods in the 'Statistical analyses' tab. The same discharge and ‘Analysis time window’ as used in the previous bicoherence analysis will be used for the following analyses.

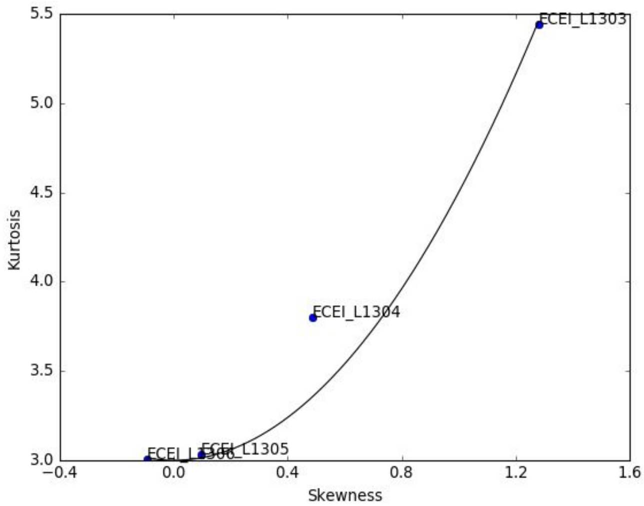
Users can check Skewness (S) and Kurtosis (K) of the target channels data by clicking the following 'PLOT' button.



The result figures show the raw data with S and K and their distribution in the SK plane. S other than 0 and K other than 3 mean a non Gaussianity of the probability distribution function of the raw data values.

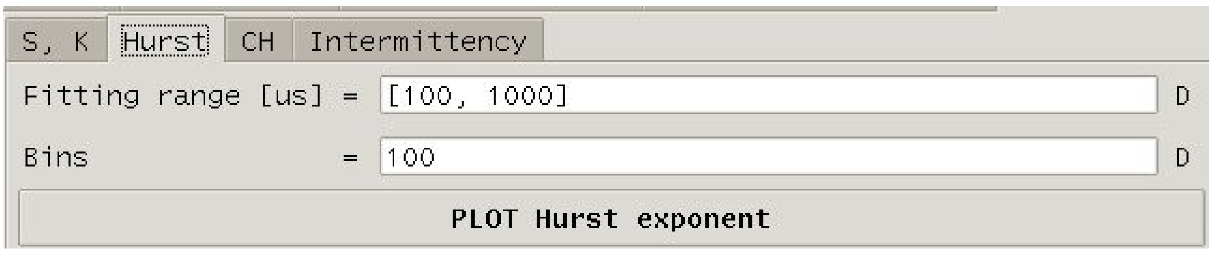


Note that ‘ECEI_L1303’ is strongly non-Gaussian as expected from the previous bicoherence analysis [4] and raw data of ‘ECEI_L1305’ and ‘ECEI_L1306’ may be dominated by the Gaussian noise.

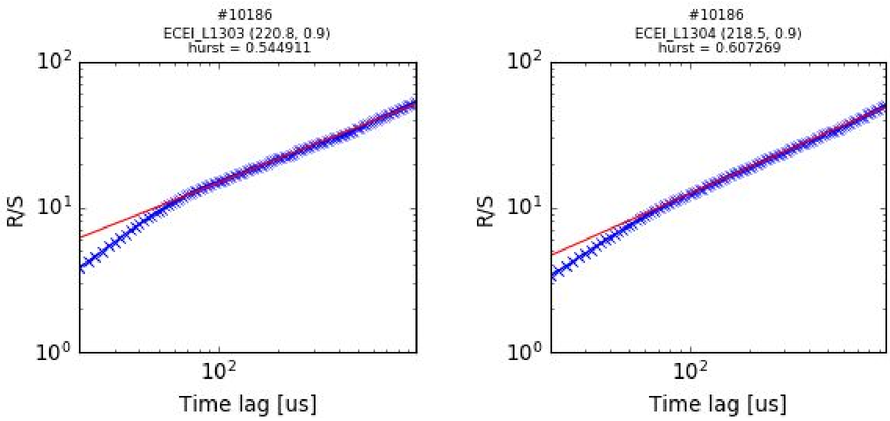


The black guide line indicates values for the Gamma distributions.

Users can check the Hurst exponent of the target channels’ data by clicking the following 'PLOT' button. The Hurst exponent will be estimated using the R/S method [5] over the 'Fitting range' of the time lag. 'Bins' could be set such that the period of 'Analysis time window' = 'Bins' x the maximum of 'Fitting range' (in our case, 100 ms = 100 * 1000 us)

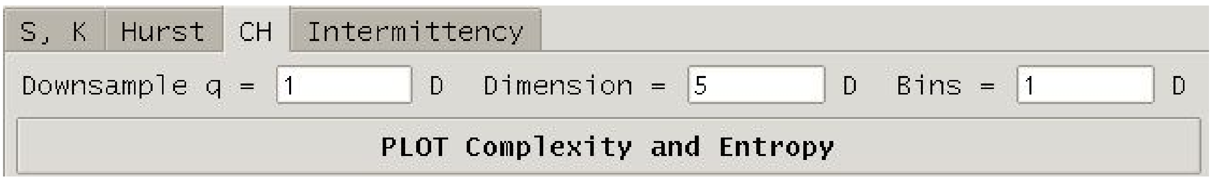


The resulting figure shows (ensemble averaged) R/S (adjusted range normalized by the standard deviation) values at different time lags and the Hurst exponent is given by the fitting exponent.

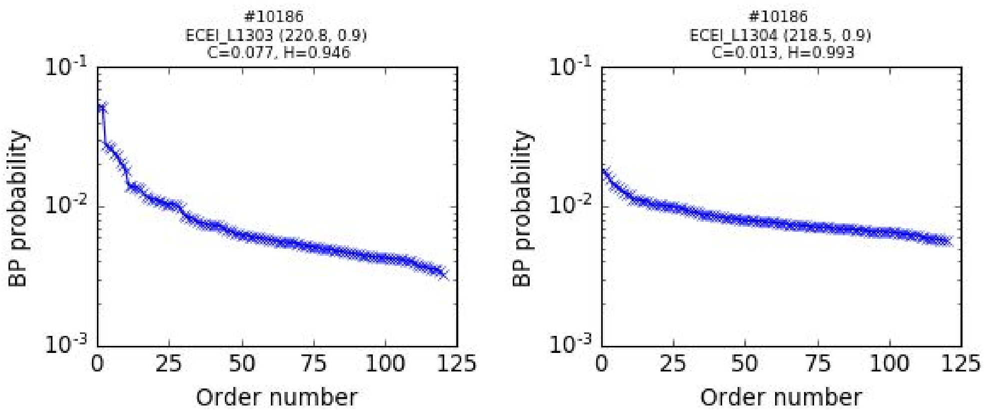


The Hurst exponent can identify a long memory of the signal. The Hurst exponent > 0.5 (< 0.5) indicates a long range temporal correlation (anti-correlation), and the Hurst exponent ~ 0.5 means no long range dependence of the signal.

Users can check the Jensen Shannon Complexity (C) and the normalized Shannon Entropy (H) of the target channels data by clicking the following 'PLOT' button. The calculation follows the method introduced by Rosso [6]. 'Downsample q' means an order of down sampling. The time step becomes q*the time step after the down sampling. 'Dimension' means the embedding dimension for the calculation of the Bandt-Pompe (BP) probability distribution function. d*the time step should be similar to the time scale of the interesting dynamics. The calculation will be done after dividing the data into a 'Bins' number of the smaller size data if 'Bins' > 1.

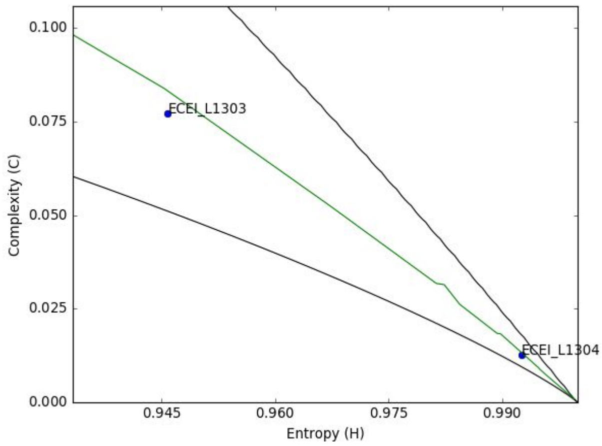


The figure shows the BP PDFs and the distribution in the CH plane.

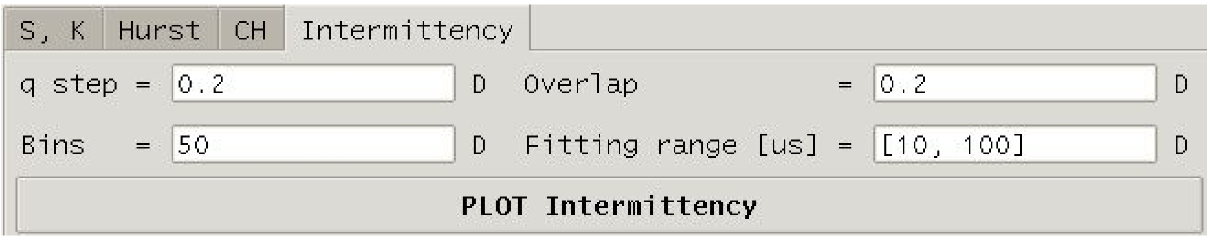


The more peaked PDF in ‘1303’ results in the higher Complexity and smaller Entropy.

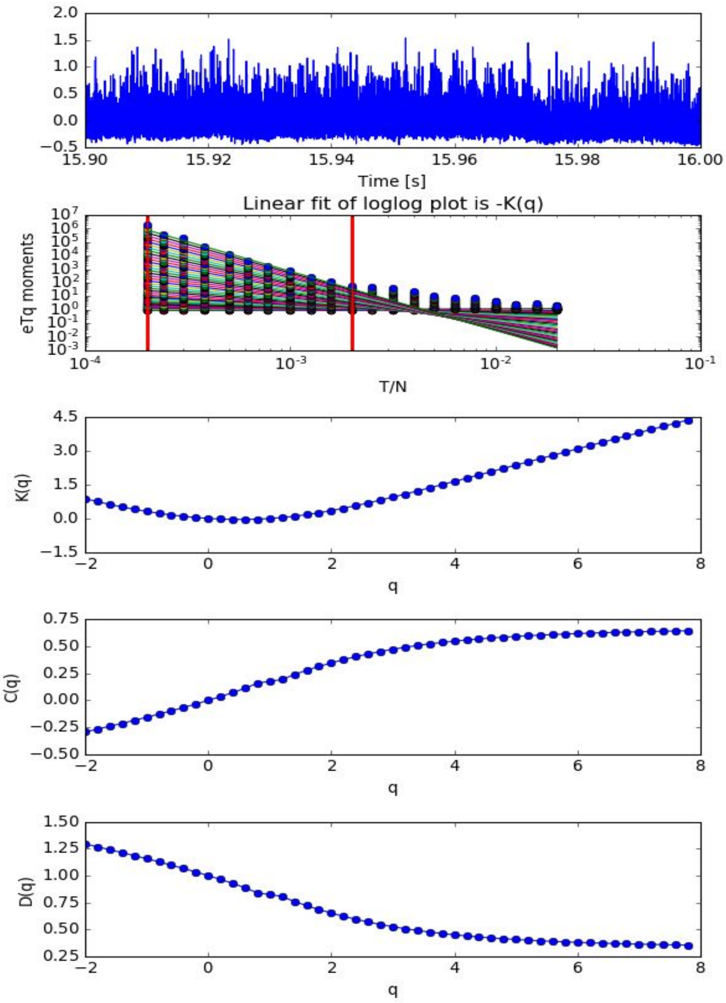
C is a measure of distance from the equilibrium (flat) PDF, and H is a measure of disorder (flatness) of the PDF. Black lines in the CH plane indicate the maximum and minimum C for each H. The moderate values of C and H above the green line indicates that the signal is more chaotic. The lower value of C and higher value of H close to or below the green line implies that the signal is more stochastic. The result implies that both are stochastic rather than chaotic, but users may need to see evolution of those values to distinguish the noise contribution.



Users can check intermittency of the target channels data by clicking the following 'PLOT' button. The calculation follows the method explained by Carreras [7]. 'q' is the moment order and ‘q step’ is the interval of calculations. 'Overlap' is the overlapping ratio between data segments, and ‘Bins’ is the number of ensembles for the calculation.



The resulting figure shows each process of the calculation from the top. Intermittency C(1) ranges between 0 (for the pure mono-fractal structure) and 1 (for the strongly multi-fractal structure).



It gives C(1), intermittency, of 0.176 for ‘ECEI_L1303’, and it is relatively higher than others. That is consistent with the higher skewness and kurtosis of this data.

Reference

A more detailed explanation and reference for each method will be provided via https://arxiv.org/abs/1907.09184.

[1] P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method based on Time Averaging Over Short, Modified Periodograms," IEEE Transactions on audio and electroacoustics, vol. AU-15, no. 2, pp. 70–73, 1967.
[2] C. Torrence and G. P. Compo, "A Practical Guide to Wavelet Analysis", Bulletin of the American Meteorological Society, vol. 79, pp. 61–78, 1998.
[3] J. M. Beall, Y. C. Kim, and E. J. Powers, "Estimation of wavenumber and frequency spectra using fixed probe pairs," Journal of Applied Physics, vol. 53, no. 6, pp. 3933–3940, 1998.
[4] Y. C. Kim and E. J. Powers, "Digital Bispectral Analysis and Its Applications to Nonlinear Wave Interactions", IEEE Transactions on Plasma Science, vol. PS-7, pp. 120–31, 1979.
[5] B. B. Mandelbrot and J. R. Wallis, "Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence", Water Resources Research, vol. 5, pp. 967–88, 1969.
[6] O. A. Rosso et al., "Distinguishing Noise from Chaos", Physical Review Letters, vol. 99, p. 154102, 2007.
[7] B. A. Carreras et al., "Intermittency of plasma edge fluctuation data: Multifractal analysis", Physics of Plasmas, vol. 7, pp. 3278–87, 2000.

Release History

The first release of the OMFIT-Fluctana version 1.0 is on June X, 2020.

The version 1.0 includes

- Spectrogram and cross spectrogram plot (short-time FFT base)
- Correlation & correlation coefficient analysis (FFT base)
- Correlation image plot
- Linear spectral analyses such as (either FFT or Morlet wavelet base)
  - Cross power spectrum
  - Cross coherence spectrum
  - Cross phase spectrum
  - Local wavenumber and frequency spectrum
- Nonlinear spectral analyses such as (either FFT or Morlet wavelet base)
  - Bicoherence
- Statistical analyses such as
  - Skewness and Kurtosis
  - The Hurst exponent
  - The Jensen-Shannon Complexity and the normalized Shannon Entropy
  - Intermittency

The version 1.0 supports only the KSTAR MDSplus data and the KSTAR ECEI data.