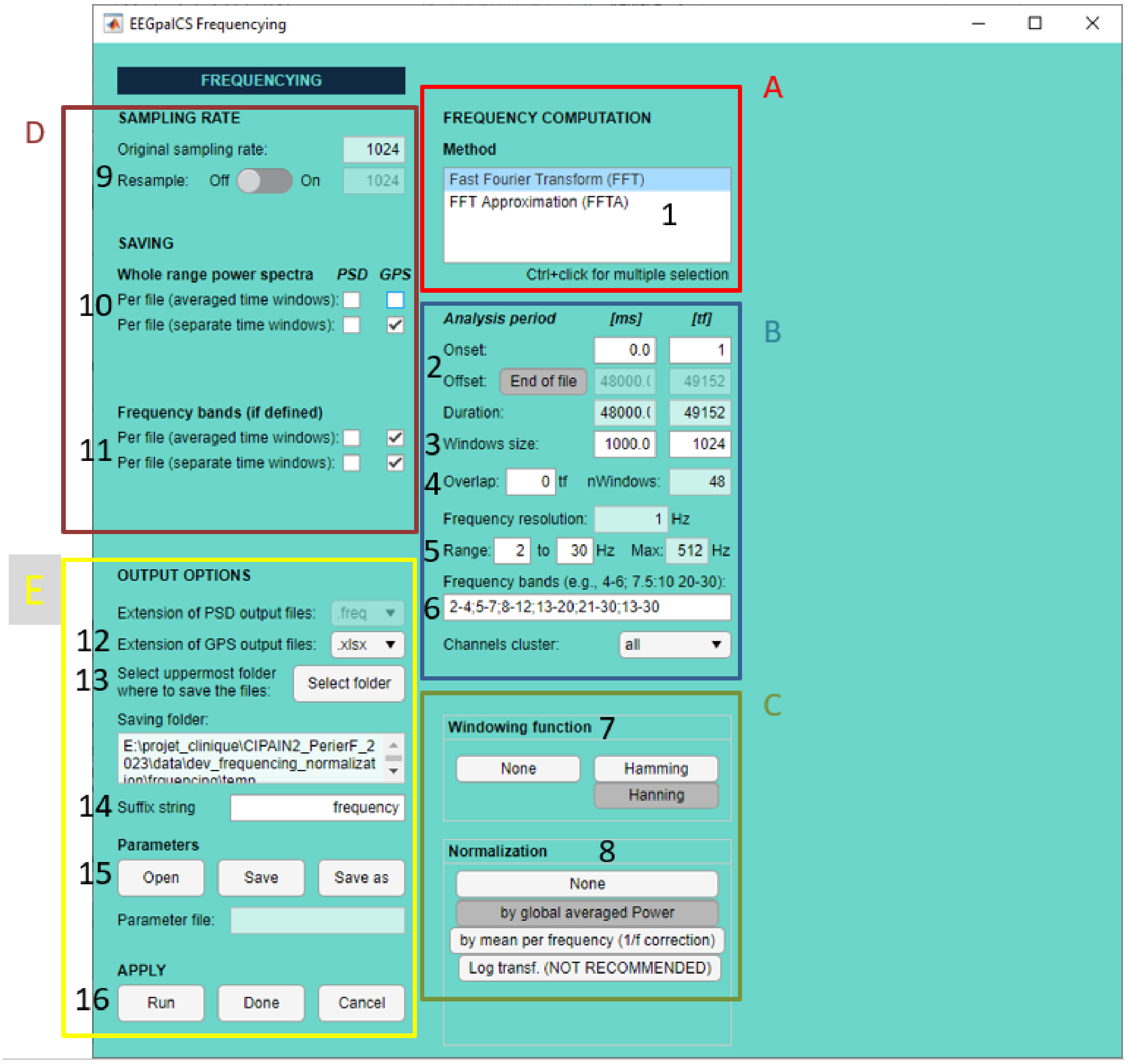
EEGpal: **Frequencying module**

Version 1.3, 10.11.2024

The 'Frequencying' module can be used to performed frequency analysis in the EEG signal. This type of analysis is often used to find biomarkers of pathologies (e.g. using a resting state recording). The output of this toolbox is the **power** for a given frequency (unit : $\mu V^2/Hz$). This power represents the amount of energy contained in these frequencies. In other words, it tells you how much these frequencies contribute to the EEG signal in the time domain. You can choose to calculate this value for each electrode separately (corresponding to the Power Spectral Density [PSD]) or to have an average/global value for your EEG signal (corresponding to the Global Power Spectrum [GPS]). This calculation is done thanks to the Discrete Fourier Transform (see FAQ for more details).



Pannel A

1. Choose the method to apply for the frequency analysis:
    a. Fast Fourier Transform (FFT) is the classical way. It provides you a power which will be positive in any case. This is enough for classical GPS analysis.
    b. FFT approximation (FFTA) is needed if you want to perform an Inverse solution (IS)analysis or topographical analysis with your frequency. More details can be found in FAQ.

Pannel B

2. Select the part of the signal on which you want to perform the frequency analysis (onset = start point and offset = end point). Reminder: [ms] = millisecond, [tf] = time frame, End of file = automatically selects the end of the file as offset.
3. **Windows**: To perform the frequency analysis, the original signal is divided into several windows. As an output, we will have a power for each window, allowing us to see the evolution over time.
   Please specify the window size (time length). Note: This parameter affects the frequencies you can study (see FAQ for more details).
4. **Overlap**: You can reduce the distance between two windows by allowing them to overlap. If you have windows of 100 TF and you have specified an overlap of 20 TF, this means that the next windows will start at 80 TF instead of 100 TF. The main purpose is to artificially increase the temporal resolution of our frequency analysis in the case of a sort recording.
5. **Frequency range to study**: This is where you specify the frequency range to be examined in your analysis. The upper limit should not be higher than the maximum frequency displayed in the GUI. The frequency resolution tells you the minimum distance between two frequencies that will be evaluated.
6. **Frequency bands**: If you are interested in specific frequency bands (e.g. alpha = 8-12 Hz, beta = 13-30 Hz), you can specify them here. Please respect the typography by using '-' to separate the upper and lower limits and ';' to separate the frequency bands (e.g. 8-12;13-30).
   Warning: The specified frequency bands must be within the frequency range specified in point 5. You cannot specify a band 13-30 if you have specified a frequency range of 0-15.

Pannel C

7. Choose the Windowing function between None, Hamming or Hanning. For more information, please read the FAQ section.
8. You can normalize the power value obtained to make the comparison between individuals or groups easier to interpret (for more detail, please consult the FAQ).
    a. **By global averaged Power**: Divide the power of each frequency by the mean power over all the frequencies defined in step 5. As result, you will obtain the distribution of frequencies relative to the mean power (for example, a value of 0.8 will means that the power at this specific frequency is 80% of the average power)
    b. **By the mean per frequency (1/f correction)**: Each frequency power in divided by the mean power of this particular frequency over all the electrodes. This has for consequence to flat the power distribution across frequencies to correct the higher power found in low versus high frequencies
    c. **Log. Transf.**: Similar as previous 1/f correction in $\log_{10}$ (similar as EEGLAB). This is not recommended due to strange result and literature (look at the FAQ for more details).

Pannel D

9. **Resample the data**: You can resample your data to another sampling frequency (new frequency must be specified in Hz). This is typically used when you want to downsample your data.
   WARNING: If you are using this option, please go through and check steps 2-6 as it will affect their values.
10. **Choose the output to save**: You can choose which power to store for the frequency range you specified in step 5. The PSD provides the frequency power for each EEG electrode and each individual frequency specified in the range. The GPS provides you with an average across these electrodes for each individual frequency specified in the range. You have to options:
    a. Per file (averaged time windows): Average the frequency power over all the time windows
    b. Per file (separate time windows): Record the frequency power for each time window. This is useful if you want to study the evolution over time.

11. The same option as described in point 10, but for the frequency band you have specified in step 6. These options are only available if you have entered a frequency band of interest.

Pannel E

12. Select the format for the output files for the GPS output (Excel of text file). You can not choose the output format for PSD file which will always be in .freq (use Catrool to open this file format)
13. Select the destination folder where the results files will be saved (note: it will automatically create a sub-folder call *Frequencying*).
14. The suffix added to the input filename to obtain the output filename
15. You can save a parameters file which will recode all the chosen options for a later processing (**save** and **save as**). You can use the button **open** to call a previous saved parameters file.
16. Click on **Run** to carry out the processing parameterized in the Frequencying module. The button **Done** will close the Frequencying module without perform the processing but keep in memory your parameters if you open again. The button **Cancel** closes the module without processing and without keep the entered parameters in memory.

FAQ

What is the discreate Fourier transform (FFT) and how can you compute the frequency power from it?

The Fourier transform (FFT) permits to transform a signal in time domain to the frequency domain. In the time domain, the amplitude of the signal is a direct measure of the variations in voltage (for example, in an EEG, in microvolts) over time. When we move on to the frequency domain using the Fourier transform, this temporal amplitude is translated into an **amplitude for each frequency**. A frequency with a large amplitude in the frequency domain indicates that it makes a large contribution to the total signal in the time domain.

When you apply the FFT to an EEG signal (or any other signal), you break it down into a sum of sinusoids of different frequencies. For each frequency, the Fourier transform gives a *complex coefficient* (for example a+b*i where a in the real component [Re(X(k))] and b is the imaginary component [Im(X(k))] that represents that frequency in the original signal. This complex coefficient has an *amplitude* (or modulus) and a *phase*. To compute this frequency amplitude, we can compute the norm of the complex vector with the following formula:

$$\text{Amplitude} = |X(k)| = \sqrt{\text{Re}(X(k))^2 + \text{Im}(X(k))^2}$$

The *power of a frequency* is a measure that describes the amount of energy contained in that particular frequency. It is calculated by taking the *square of the amplitude* of that frequency (i.e. the square of the magnitude of the coefficient in the frequency domain).

Formellement :

$$\text{Puissance de la fréquence} = |X(f)|^2$$

où $X(f)$ est le coefficient complexe pour la fréquence f dans la transformée de Fourier, et $|X(f)|$ est son module (ou amplitude).

The last step is to divide this power by the size of temporal segments (N) to obtain the *average power per unit time* for each frequency. This operation normalizes the power, making it comparable between different epoch lengths or between different segments of the signal.

$$\text{Puissance de chaque fréquence} = \frac{|X(f)|^2}{N}$$

What is the difference between the FFT and FFTA and when should I use the FFTA ?

The problem with the classical FFT is that you lose the polarity information needed to compute an IS. This is why Brunet et al. (paper: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3022183/) proposed a way to restore the polarity information between the electrodes. It is the famous FFTA introduced in the Cartool software. In fact, the reconstructed polarity is random and an electrode can be positive on one map and negative on another. To solve this problem, Cartool will use the polarity of the first map to correct all the others (to have the same polarity everywhere). MichaelDP improved the concept by looking for the best map to use as a reference by studying the maps that correlate the most.

Why does the choice of window length affect my frequency analysis (point 3)?

The highest frequency you can analyse is always half the window length. For example, if my window is 1000 ms, the maximum frequency of my analysis will be 500 Hz (in the case of a sampling rate at 1000Hz). This is because of the Nyquist theorem. The maximum frequency is shown in the interface at position 5.

What is the windowing function and why should I use it (point 7)?

During frequency analysis in EEG, we cut the signal in several time window which introduce discontinuities. The use of a *windowing* function may be necessary to reduce the undesirable effects associated with discontinuities at the edges of the signal when calculating the Fourier transform. These discontinuities manifest themselves as ‘edge artefacts’ that create undesirable effects in the power spectrum, in particular the phenomenon of '*spectral leakage'*, where the energy of a given frequency 'leaks' towards adjacent frequencies.

Spectral leakage occurs mainly because of discontinuities at the ends of a signal segment. When performing a Fourier analysis on a finite time segment, it is implicitly assumed that the segment is periodic (that it repeats ad infinitum). If the start and end values of the segment do not match, a discontinuity is created, introducing high frequencies that are not present in the original signal.The consequence could be :

- Dispersion of energy over several frequencies: Instead of being concentrated on the frequency of interest, the energy 'leaks' into neighbouring frequencies, making the peak less clear and making it difficult to identify certain frequencies accurately.
- Difficulty in detecting nearby frequencies: In an EEG signal, where you want to analyse narrow frequency bands (alpha, beta, etc.), spectral leakage can mask details and make it difficult to distinguish between nearby frequencies.
- Bias in the power spectral density (PSD): As the energy is not strictly localised on the expected frequencies, the measurement of the power of the frequency bands is less reliable.

The windows function search as Hamming or Hanning will modulate the amplitude of the signal in the time domain (before FFT computation) to decrease the contribution of windows edge. For example, if I have a 1024 TF window, we'll multiply the data by a vector [0.08, 0.08,...,1,...,0.08) for Hamming windowing.

The disadvantage of applying a windowing function is that it artificially reduces the amplitude at the ends of the signal, which can reduce the power measured in the frequency bands, as it attenuates the signal in the time domain before moving on to the frequency domain. To compensate this lost of power, you can normalize the powers according to a reference period or a reference frequency as suggested in point 8.

Sould I use a hamming or an hanning window function (point 7)?

For EEG frequency analysis in the 2-30 Hz spectrum, the Hamming and Hanning windows are both good options for reducing spectral leakage, but they have slightly different characteristics that can guide the choice according to the precise objectives of the analysis:

1. *Hanning window:*
    - This window offers progressive attenuation at the ends of the segment, making it very effective at limiting spectral leakage.
    - It has better control of near frequencies, but with a wider central lobe than the Hamming window.
    - Hanning is particularly suitable if the aim is to isolate relatively narrow frequency bands accurately, such as the alpha and beta bands in EEG, while minimising leakage.
2. *Hamming window:*
    - This has a slightly different shape and offers a better compromise between leakage attenuation and central lobe width, which improves frequency resolution (i.e. the ability to discriminate frequency peaks).

- It attenuates the edges less, which reduces the loss of information at the edges and can be an advantage if you have a long-duration signal.
- Hamming is often preferred if you're looking to balance leakage and frequency resolution, especially when you're interested in slightly wider frequency bands.

Conclusion

For a broad spectrum such as 2-30 Hz in EEG:

- *Hanning window*: preferred if you are looking to minimise spectral leakage, particularly for narrow bands.
- *Hamming window*: recommended if frequency resolution is important, with a slight compromise on leakage.

In practice, the two windows give similar results for EEG analysis, but if you are working with very close bands (for example, a fine separation between alpha and beta), the Hanning window may be slightly more appropriate.

Text written with the help of ChatGPT.

Should I normalize my power values (point 8)?

Whether or not to normalize power values depends on your specific goals and the type of analysis you're performing. Here are some cases where normalization is often recommended and how it can improve the interpretability of your power results:

Comparing Across Participants or Trials

- *Why Normalize?* Individual differences (e.g., baseline power or amplitude differences between participants) can make it difficult to compare power values across participants or sessions.
- *How to Normalize?* You can normalize power within each participant, often by dividing by the total power (option *by global averaged Power* in point 8)

Comparing Power Across Different Frequency Bands

- *Why Normalize?* Power values naturally tend to be higher in lower frequency bands due to the 1/f property of EEG signals (where lower frequencies have higher power). Normalizing helps account for this, especially if you're interested in comparing higher-frequency bands (e.g., alpha or beta) to lower-frequency bands (e.g., delta or theta).
- *How to Normalize?* Some approaches include normalizing each frequency band by the average or baseline power in the entire signal or by using a 1/f fit to flatten the power spectrum (option *by mean per frequency (1/f correction)* in point 8.

In addition, Log-transforming power values (e.g., log10) can reduce the skew from extreme values and is often used in EEG power spectral analysis to make distributions more normal. The formula is

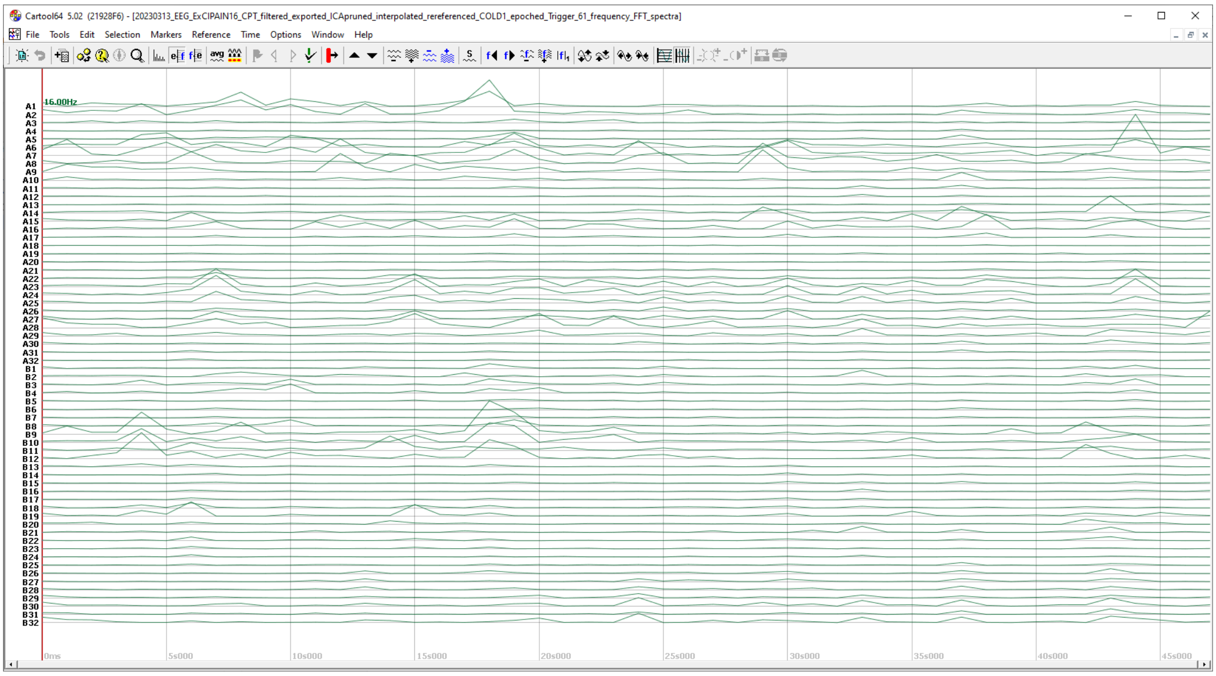
LogPower = 10 .* $\log_{10}$(Power ./ mean power);

However, this normalization has been criticized in the paper: https://pmc.ncbi.nlm.nih.gov/articles/PMC8354524/

Text written with the help of ChatGPT.

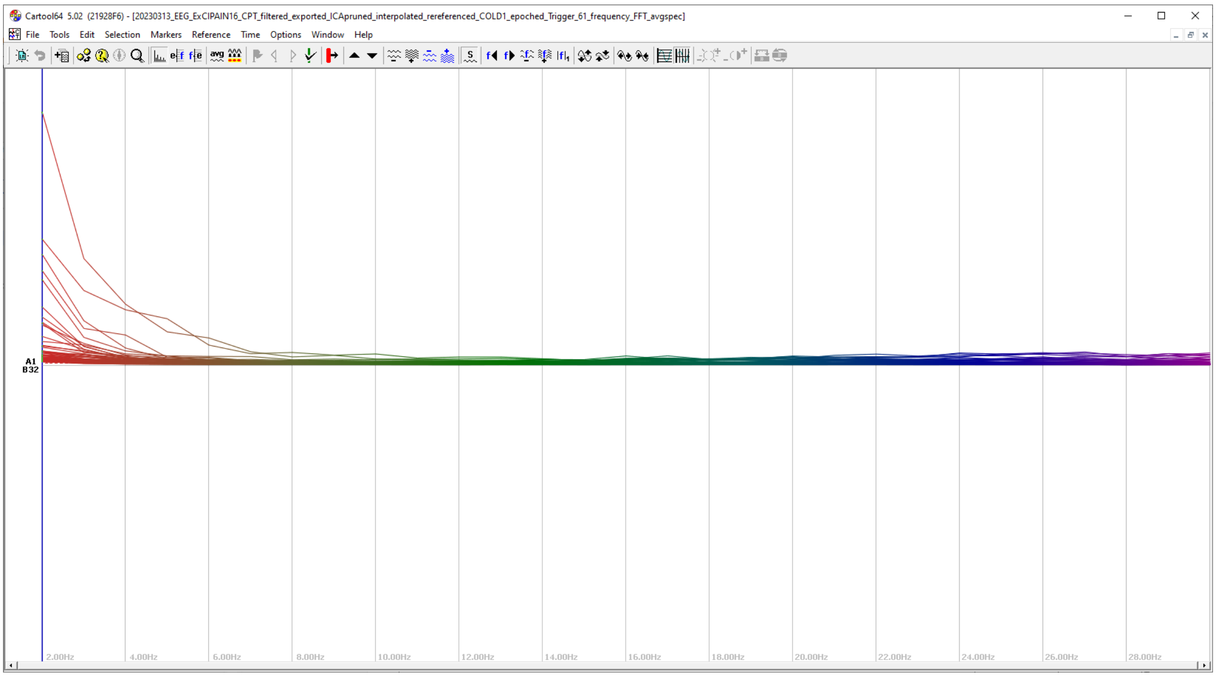
How can I display my result .freq file in Cartool?

In the case of 'Per file (separate time windows)' option from point 11



By default, Cartool show you the result for a specific frequency (in this example 16Hz) with the power for each time windows (horizontal dimension) and the electrodes (vercital dimension). To change the frequency to displayed, use the button : 

In case of 'Per file (averaged time windows)' option :



By default, Cartool show you the Power spectrum view. Each line is the average power across each time window for each electrode. The horizontal axis is the different frequency of the spectrum.

In both cases, you can associate the file with a coordinate file to have an estimation of the topography (in a lm file) as usual. The manipulation of the power trace is also similar as in EEG file display.