Tutorial 5

Probability Density

The **Probability Density** module includes four analysis tools, all related to the Probability Distribution Functions (PDFs):

1. **PDFs** computes the probability distribution functions for a range of scales;
2. **Flatness** computes the parameter with the same name related to the fourth order moment of the PDFs;
3. **OPR** computes a parameter to be used in an attempt to rescale all the computed PDFs;
4. **OPR Rescaling** applies the One Parameter Rescaling (OPR) procedure based on the parameter computed by OPR.

An example of the initial window of this module is given in Figure 1.
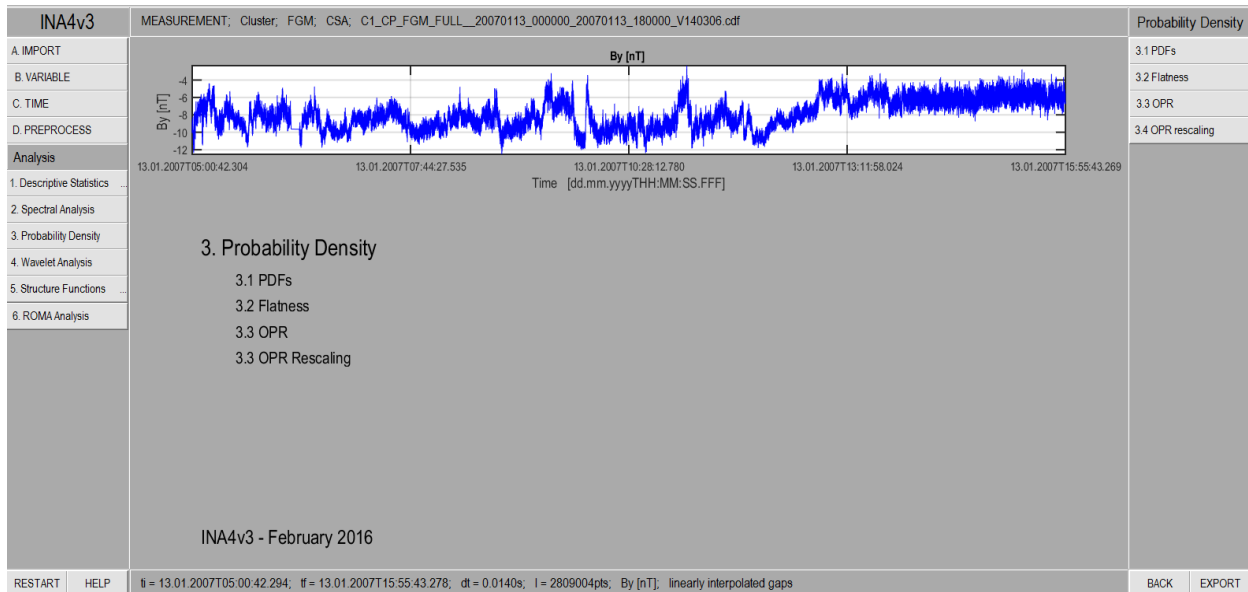

Figure 1. The main window of PDFs analysis opens when the user clicks the **Probability Density** button of the left hand side panel of INA. The upper plot shows the time series to be analyzed. The four buttons at the right are links to the analysis routines devoted to PDFs analysis: the computations of the PDFs themselves, the computation of the flatness parameter, the computation of the one parameter rescaling (OPR) and the rescaling of the computed PDFs with the OPR parameter.

Each tool is illustrated and described below.

a) 3.1 PDFs

When the user clicks the button *3.1. PDFs* of the main **Probability Density** window shown in figure 1, a new window opens, illustrated in Figure 2. The PDFs subclass computes by default the unscaled Probability Density Functions (PDFs) for a number of N scales; each scale comprises a number of points equal to 2^tau, where tau takes values between tau_min, in general equal to 0, and tau_max-1, where tau_max is the smallest power of 2 for which 2^tau_max is still larger than the total length of the time series. The PDFs are obtained by moving a sliding window of length 2^tau over the entire time series to be analyzed and taking the difference between the right and left edge points of the window. The window is displaced by one point at each step, thus consecutive windows overlap. The normalized histogram of the differences/increments gives the PDF at that scale.

The computation of PDFs has four adjustable parameters: *tau_min*, *dtau* and *tau_max*, which define the scale range tau, and *bins* that defines the number of bins used by the histogram function to compute the increments (the default is 100). The user can choose to compute either multiple scales (tau_min < tau_max) or single scales (tau_min = tau_max and dtau = 1).
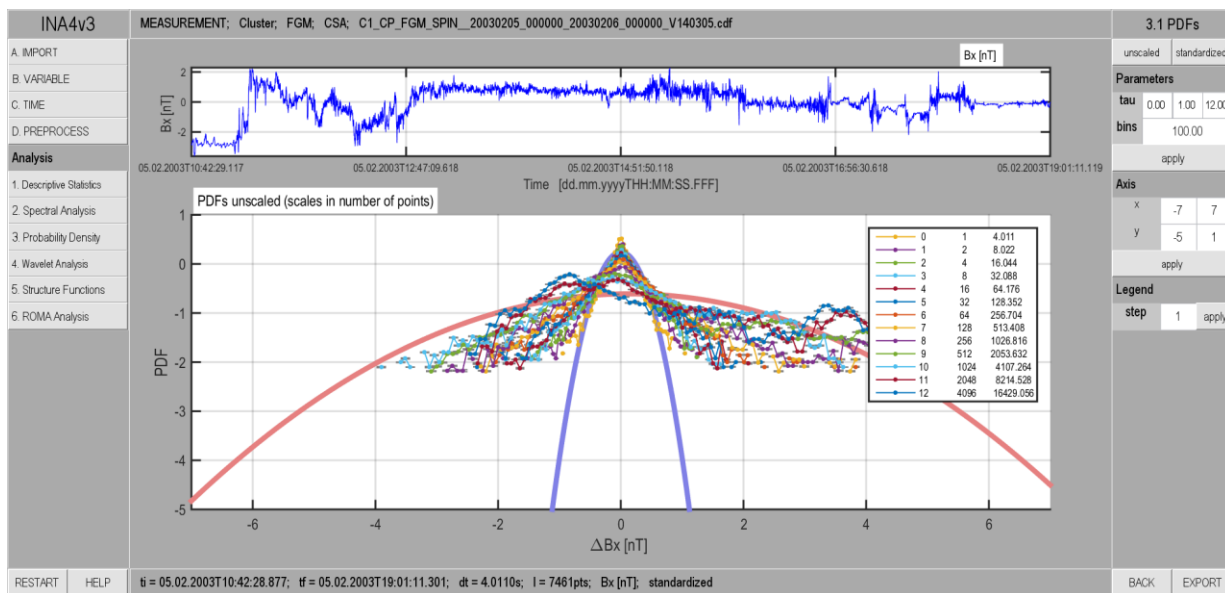

*Figure 2 Illustration of the INA window that opens when the user clicks the **3.1 PDFs** button of the main **Probability Density** window (shown in Figure 1). By default the PDFs are computed for all available scales, as shown by the central plot. The PDFs corresponding to different scales are displayed with different colors. The caption near the PDFs plot indicates the considered scales and specifies them as number of points or seconds. The user can further process the PDFs by using the control panel at the right of the figure; the field **tau** controls the minimum and maximum scales and how many scales between the two are analyzed; the **bins** field controls the number of bins used to compute the histogram; **Axis** controls the Oy and Ox limits of the PDFs plot. The **Legend** field controls how many scales are included in the caption attached to the right of the PDF plot.*

Figure 2 illustrates the PDFs subclass for the same time series discussed in the previous tutorials. The central plot illustrated the superposed PDFs of increments for 13 scales: tau ranging from 2^0 = 1 point to 2^12 = 4096 points. All the twelve intermediate scales are shown. The caption in the top right corner shows: the powers of two (first column), the number of points of the sliding window (second column) and the corresponding length of the window in seconds (third column). Two Gaussian PDFs are superposed on the computed PDFs: the blue/red Gaussian is defined such that it has the same mean and standard deviation as the PDF of the first/last scale.

The "standardized" button applies a standardization procedure on each incremental time series by subtracting the mean and dividing by the standard deviation, similarly with the PREPROCESS procedure described in Tutorial #2. Figure 3 shows the results of the standardization applied on the PDFs for the 13 scales shown in Figure 2.
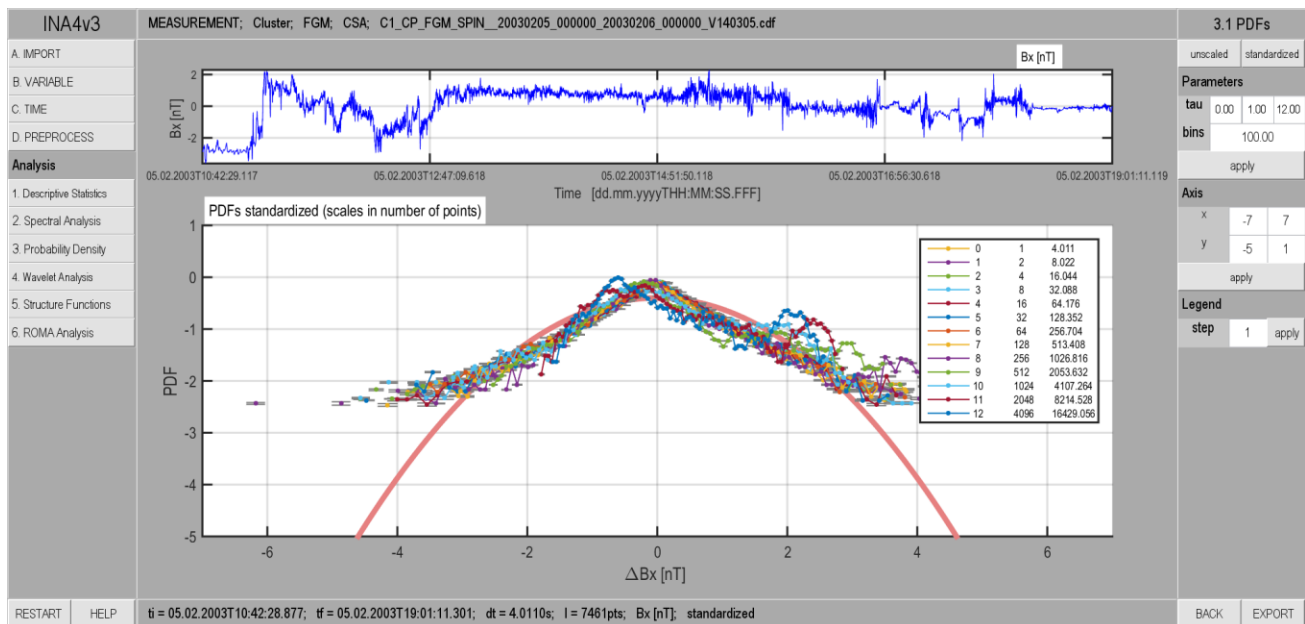

Figure 3 Illustration of the standardization procedure applied on the PDFs illustrated by Figure 2.

There are two display parameters that can be adjusted by the user:
1. **Axis** controls the minimum and maximum values of the x and y axis, and
2. **Legend** controls the number of entries displayed in the legend.

In Figures 2 and 3 the x axis goes from -7 to 7, and the y axis (given in powers of 10) goes from -5 to 1; the user can easily change the corresponding editable fields located in the control panel. The default value for the display parameter labeled **Legend** is 1, i.e., the user will see all the scales in the legend of the PDFs plot. By choosing a value of 2, the legend will show only half of the plotted PDFs, skipping every other entry, and so on. This parameter can be used

to decrease the size of the legend, which is especially useful when there are a large number of PDFs plotted on the screen.

b) 3.2. Flatness

The Flatness routine computes by default the Flatness factor:

$$F = \frac{\langle \Delta P(t,\tau)^4 \rangle}{(\langle \Delta P(t,\tau)^2 \rangle)^2}$$

where ΔP identifies the differences computed with the procedure described and $<>$ means time average, considered equivalent to ensemble averaging. The flatness is computed for the same scales τ as the ones used for computing the PDFs. An illustration is shown in Figure 4 where the red line corresponds to the Flatness of Gaussian fluctuations, equal to 3 for all scales. The Flatness computation has the same adjustable parameters as the PDFs method described above.
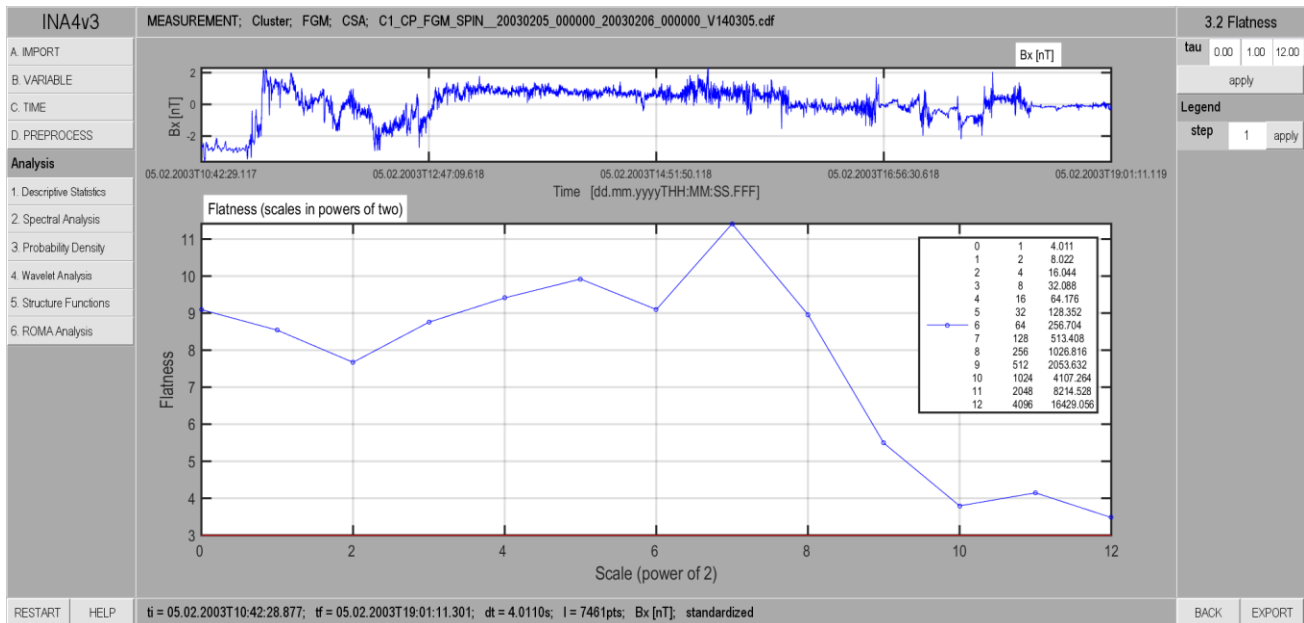

Figure 4. Illustration of the Flatness analysis window opened from the right panel of the general Probability Density analysis window (see Figure 1).

c) 3.3. One Parameter Rescaling (OPR)

This functionality aims to find a parameter to be used to rescale the PDFs. The assumption is that the distributions can be described by a stable, symmetric shape (like a Gaussian or Levy) PDF. The scaling exponent is determined from the slope of the peak value of the PDFs, P (0, tau), versus the scale. The OPR function is used to obtain P (0, tau) as a function of tau and the slope is computed automatically. Note that increasing or decreasing the number of bins used to compute the PDFs can increase the accuracy of estimating the peaks, so **bins**, is an adjustable parameter here.

d) 3.4. OPR Rescaling

A more advanced OPR implementation can be found in the **OPR Rescaling** panel, illustrated by Figure 5. The panels show an estimate of the scaling exponent s (in the plot labeled "One Parameter Rescaling") and an actual rescaling of the PDFs according to the one parameter rescaling: $P(\delta y, \tau) \cdot \tau^s = P_s(\delta y / \tau^s)$, where P_s is the scaling function onto which all the other PDFs must collapse. **s** is a variable parameter here (labeled **s_opr**), and can be set by the user to any value between 0 and 1. The plot labeled OPR rescaled, located in the bottom left corner, shows the peak value vs. scale for the rescaled PDFs.
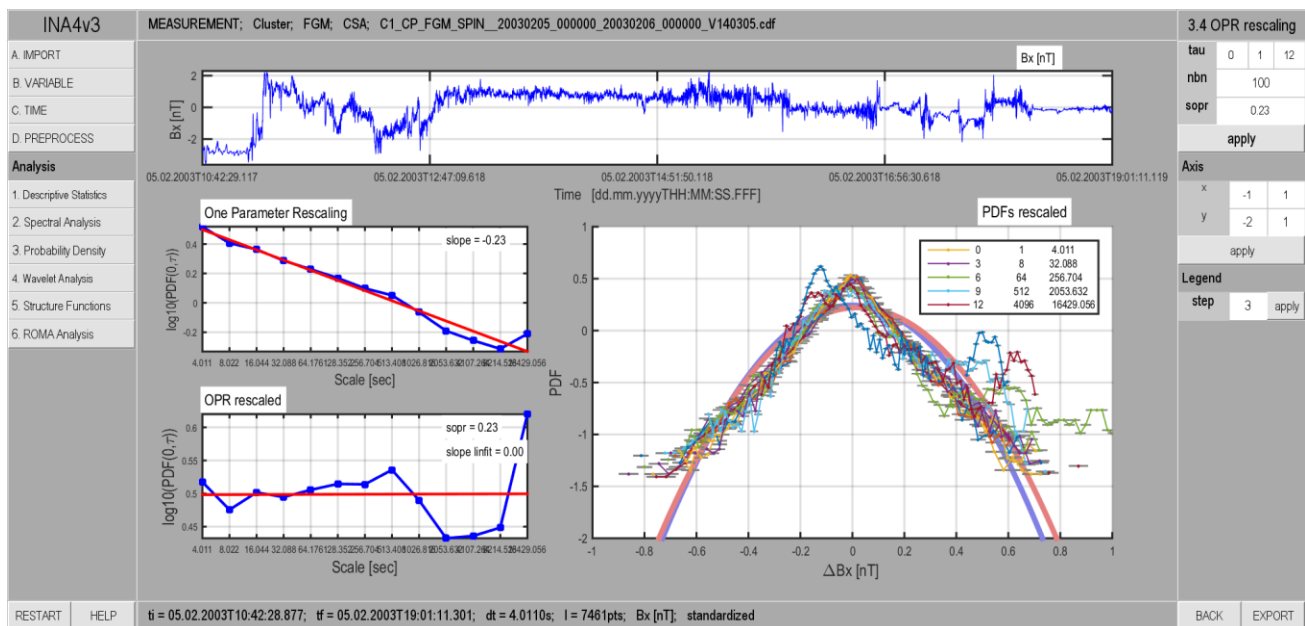

Figure 5. Illustration of the One Parameter Rescaling procedure. The upper plot shows the analyzed signal. One Parameter Rescaling plot shows how the maximum of the PDFs, P (0, τ), vary with the scale τ , and estimates the slope that determines the parameter for rescaling. OPR rescaled plot shows the rescaling of the maxima with the parameter determined above. The right plot shows how the one parameter rescaling is applied on the entire PDFs.