BRiSTOL: a Band-limited RMS Stationarity Test Tool for Gravitational Wave Data

Francesco Di Renzo,[1,2] Francesco Fidecaro,[1,2] Massimiliano Razzano,[1,2] and Nunziato Sorrentino,[1,2]

[1] *Università di Pisa, I-56127 Pisa, Italy*
[2] *INFN, Sezione di Pisa, I-56127 Pisa, Italy*

Overview. Common techniques in Gravitational Wave data analysis assume, to some extent, the stationarity and Gaussianity of the detector noise. These assumptions are not always satisfied because of the presence of short duration transients, namely *glitches*, and other slower variations in the statistical properties of the noise, which might be related to malfunctioning subsystems. We present here a new technique to test the stationarity hypothesis with minimal assumptions on the data, exploiting the band-limited root mean square and the two-samples Kolmogorov-Smirnov test. The outcome is a time-frequency map showing where the hypothesis should be rejected. This technique was used as part of the event validation procedure for assessing the quality of the LIGO and Virgo data during O3.

Strong-sense stationary process: its statistical properties are left unchanged when shifted in time (index set):

$$F_X(x_{t_1}, x_{t_2}, x_{t_3}, \dots) = F_X(x_{t_1+\tau}, x_{t_2+\tau}, x_{t_3+\tau}, \dots) \quad \forall\ \tau, t_1, t_2, t_3, \dots$$

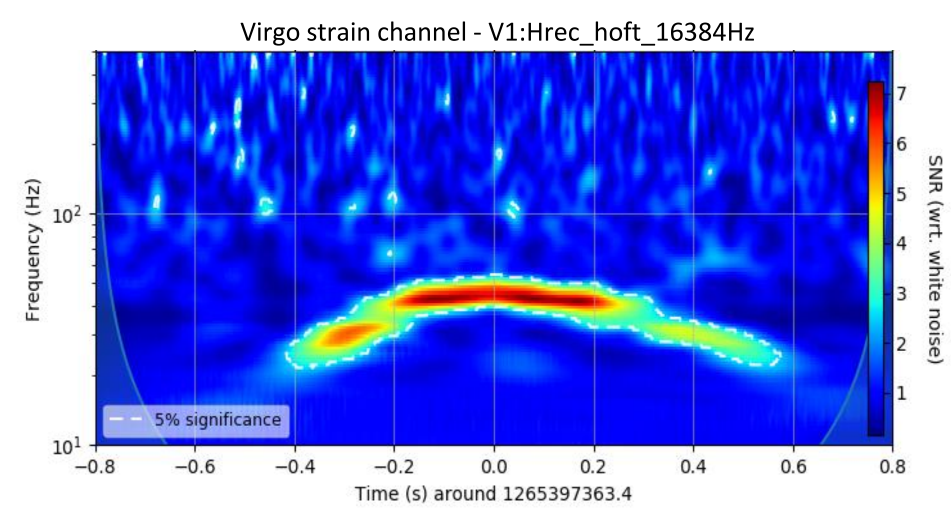
Wide-sense stationary process: first and second order moments are time-shift invariant:

$$E[X_t] = E[X_{t+\tau}] = \mu, \qquad R_X(t, t+\tau) := E[X_t X_{t+\tau}] = R_X(\tau)$$

Non-stationary noise in Gravitational Wave detectors

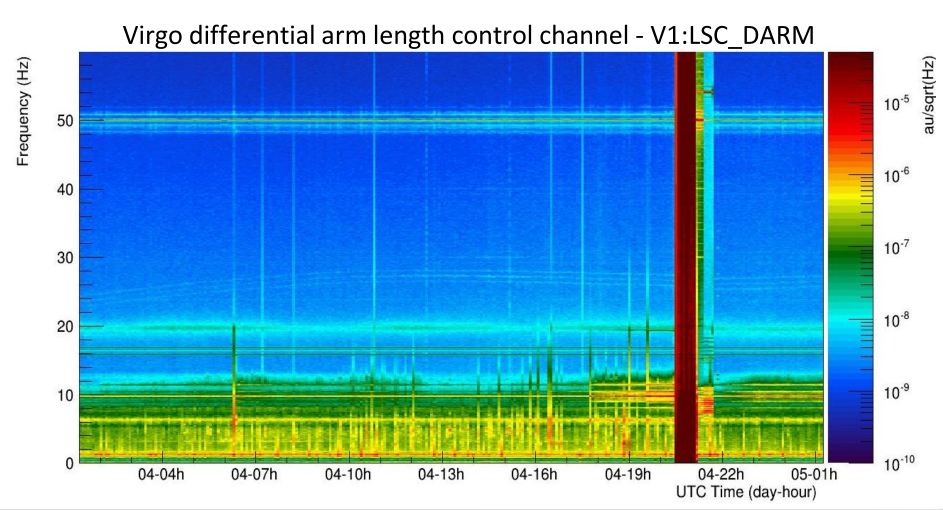
Glitches: *rapid* transient noise excesses:

- About 10 per minute with SNR > 6.5 [1];
- They can mimic astrophysical signals or mask them;
- Identified by means of Event Trigger Generators (e.g. Omicron [2]).

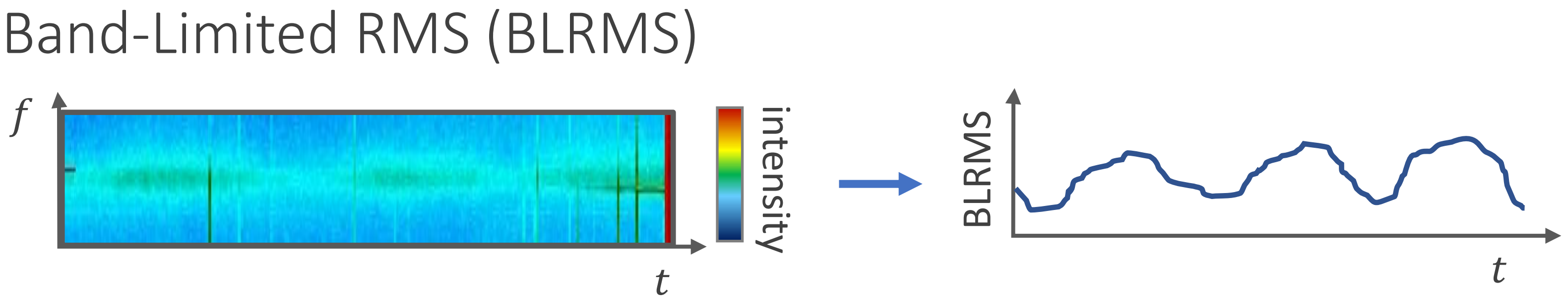


Slow non-stationarities,

- including trends, bumps, seasonal effects and wandering lines;
- Usually identified by means of spectrograms, median-normalised spectrograms, or *Rayleigh Gaussianity test*.



Band-Limited RMS (BLRMS)



$$BLRMS(t; [f_1, f_2]) = \sqrt{\int_{f_1}^{f_2} \hat{S}(t, f)\, df}$$

where $\hat{S}(f, t)$ is an estimate of the process Power Spectral Density (PSD) referred to time t, obtained e.g. via Welch's method [3].

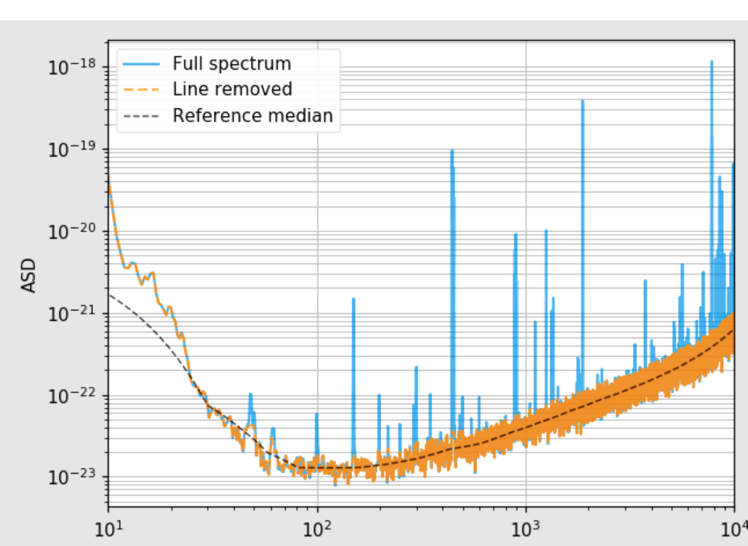
Properties:

- Dividing the spectrum in bands allows to focus on regions where *known noise sources* are likely to manifest. E.g.: [10,40] Hz scattered light, [145,155] Hz non-linear couplings of seismic motion to angular controls;
- With respect to working directly with $\hat{S}(f, t)$, averaging on frequency reduces the variance ⇒ better time resolution (like Welch's method on time [3]).

Modified BLRMS algorithm [4]

Spectral line removal, to focus on variations of the "noise floor":

- Reference median PSD estimate in some generic bands;
- Median Absolute Deviation (MAD) to identify lines;
- Interpolate and iterate.

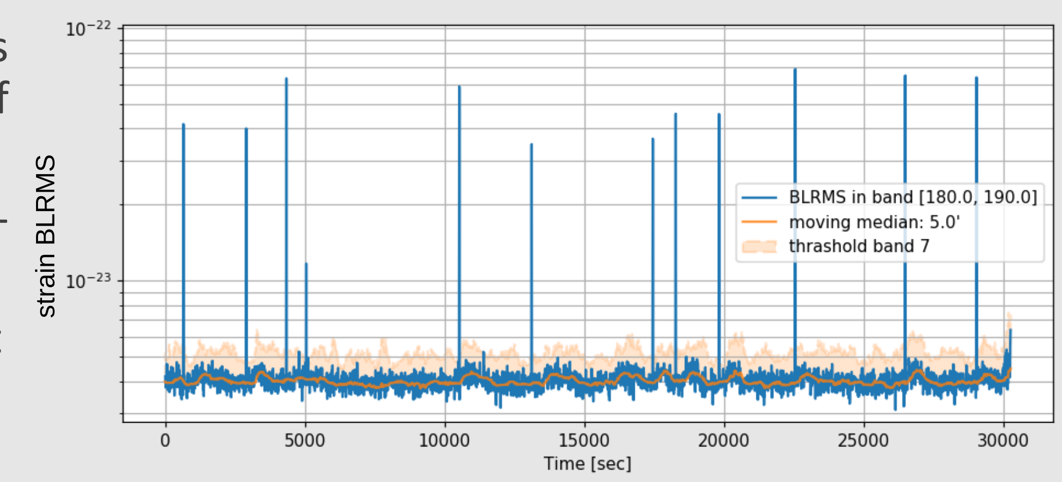


[1] Abbott R et al. 2020 (Preprint 2010.14527)
[2] Robinet F et al. 2020 (Preprint 2007.11374)
[3] Welch P 1967 IEEE Transactions on Audio and Electroacoustics 15 70-73 doi:10.1109/TAU.1967.1161901
[4] DiRenzo F 2019, VIR-0511A-19
[5] Kolmogorov A (1933). G. Ist. Ital. Attuari. 4: 83–91.

Modified BLRMS algorithm [4] (continued)

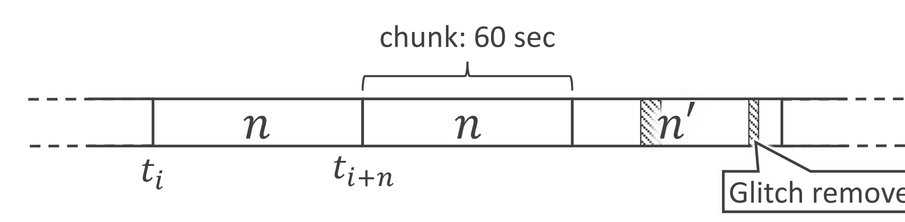
Glitch removal, to focus on "slower variations" of the background noise:

- Moving median algorithm to produce MAD;
- Adaptive threshold: last points weight more.



Stationarity test of the BLRMS empirical distribution function

1. We divide $BLRMS(t_i; b)$, $b \in \{[f_1^{min}, f_1^{max}], \dots, [f_k^{min}, f_k^{max}]\}$, into chunks:

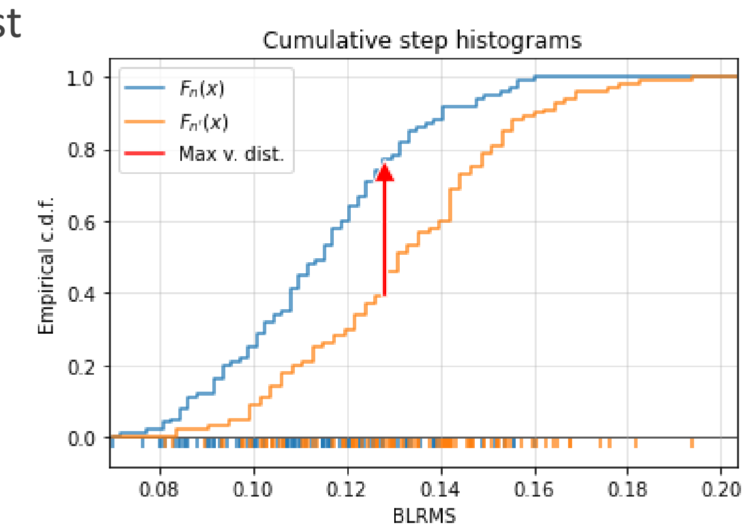


2. "Two samples" Kolmogorov Smirnov test [5] on the empirical cumulative distribution function of the BLRMS',

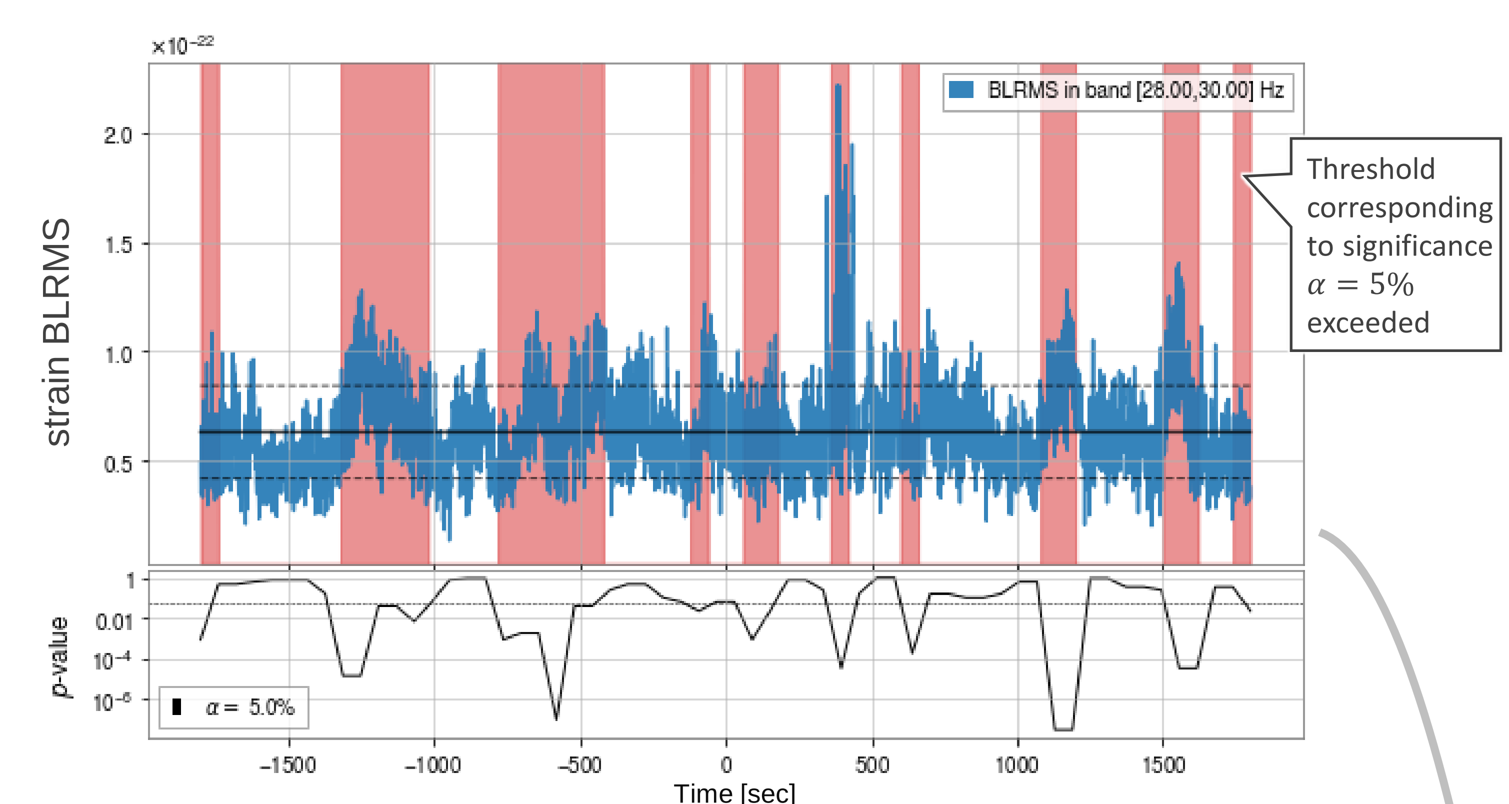
$$F_n(x) = \sum_{i=1}^{n} I(X_i \leq x)$$

of the data in neighbouring chunks:

$$KS_{nn'} = \sqrt{\frac{nn'}{n+n'}} \sup_x |F_n(x) - F_{n'}(x)|$$

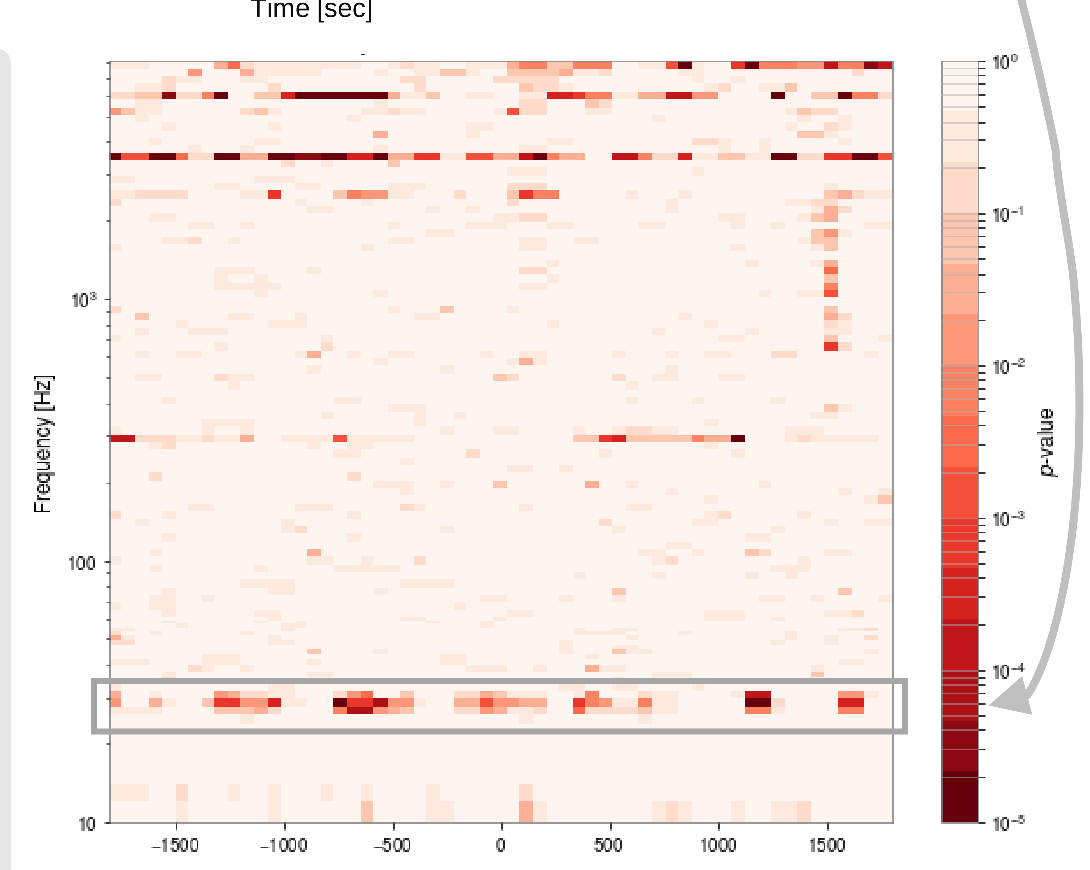


3. Stationarity (null hypothesis, H_0) rejected with **significance level α** on the base of the **p-value** of the previous test statistic (Kolmogorov distribution).



Notes

- Different possibilities for the metric choice: Kramer-von Mises (L^2), Anderson-Darling, Cucconi;
- Trade-off in the sensitivity to central tendency, dispersion, and tails of the c.d.f.;
- It allows to test Gaussianity independently;
- Part of Virgo DQR since O3.



55th Rencontres de Moriond 2021, Gravitation session

francesco.direnzo@pi.infn.it