Coherent WaveBurst update for LHV network.

S.Klimenko, G. Vedovato

LIGO note T-xxxxxx

1. Adaptive regulator

The general purpose of the CWB regulator is to constrain the x-component of the signal when $|f_+| >> |f_x|$ and the x-component is completely burried in the detector noise - this is very tipical for existing detector networks (particularly the LH network). The detectability of the x-component depends on the detector netwiork and the signal SNR, therefore the regulator needs to be carefully selected. The purpose of the adaptive regulator is to avoid the fine tuning of the analysis for different detector networks and improve detection of very high SNR events (SNR¿100). It also solves the following problem: given a network of several detectors, not all detectors are participating equally for measurements of the GW signal. This is particularly true when some detectors are much less sensitive than others, like Virgo detector in the O2 run. More important, the relative contribution of each detector varies during the run and depends on a) position in the sky and b) frequency band of detected signals. Namely, some detectors may be only spectators adding noise to the network and the individual detector contribution varies from signal to signal. It is not possible to adjust regulator for each event, therefore the adaptive regulator is required. The details of the adaptive regulator construction are described here [1]

2. Statistics of the residual noise

Residual noise is defined as $r = x - s$, when signal **s** is subtracted from the data **x**. The statistics of the residual noise should follow the ξ^2 distribution. The ξ^2 statistic is defined as $\xi^2 = E_n/N_{DoF}$ where E_n is the noise energy and N_{DoF} is the number of degrees of freedom. The noise energy is estimated as sum of two terms $E_n = E_r + E_g$, where $E_r = \int(x(t) - s(t))^2 dt$ is the energy of the residual noise integrated over the event and E_g is the estimated contribution of the Gaussian noise remaining in the signal. The ξ^2 statistic can be calculated in the TF domain and time domain. Significant deviation of ξ^2 from unity is an indication of a glitch. In the TF domain the Er is equal of the sum over all pixels $(i = 1, .., I)$ in the TF cluster: $E_r = \sum_i (w[i] - s[i])^2 + (W[i] - S[i])^2/2$, where $w[i], W[i]$ are the 0-phase,90-phase WDM amplitudes and $s[i], S[i]$ is the reconstructed signal. The estimator of the Gaussian noise E_g absorbed in the signal depends on the regulator: $E_g = \sum_i G_n[i]$, where

$$G_n[i] = (|f_+[i]|/f[i])^2 + |f_\times[i]|/F[i])^2. \quad (1)$$

The regulated antenna pattern amplitudes $f \geq |f_+|$ and $F \geq |f_\times|$ are defined here https://wiki.ligo.org/pub/Bursts/Cwb20170518/regulator.pdf When no regulator is imposed $E_g = 2I$, for hard regulator $E_g = I$, $N_{DoF} = K * I$, where K is the number of detectors. In the time domain the residual energy is $Er \int (x(t) - h(t))^2 dt$, where

$x(t)$ is the time domain representation of the cluster data and $h(t)$ the reconstructed signal obtained with the WDM packets. In general both $x(t)$ and $h(t)$ are weighted sums over the WDM basis functions weighted by the pixel norms $n[i]$ calculated from the cross-talk coefficients with the neighbour pixels [2]. The fundamental property on the pixel norms $n[i]$ is that the $\sum_i 1/n[i]$ gives the number of degrees of freedom in the time domain: $N_{DoF} = \sum_i 1/n[i]$. The estimated contribution of the Gaussian noise in the time domain is calculated as

$$E_g = \sum_i G_n[i]/n[i]. \tag{2}$$

Global norm of the event: Depending on the event SNR, the data in the TF domain is oversampled by representing the event at several TF resolutions. As a result the event TF energy $E_{tf} = \sum_i w[i]^2 + W[i]/2$ and time-domin energy $E_t = \int x(t)^2 dt$ could be quite different. The oversampling factor is described by the event $N_{norm} = E_{tf}/E_t$, which has a meaning of the average number of WDM resolutions used for the event reconsreuction. The low SNR events are reconstructed with $N_{norm} \sim 1$. Very high SNR events are reconstructed with the global norm is equal to the number of WDM resolutions.

3. Coherent energy

The coherent energy E_c is calculated in the TF domain as a sum over coherent energies of each pixel ($E_c = \sum_i E_c[i]$), where

$$E_c[i] = |\mathbf{s}[i]|^2[1 - \frac{\sum_k s[k,i]^4}{|\mathbf{s}[i]|^4}] + 90^o term, \tag{3}$$

where s[k,i] are the detector response amplitudes - components of the vector **s**[i]

$$E_c[i] = |\mathbf{s}[i]|^2 * [1 - \frac{\sum_k s[k,i]^2 w[k,i]^2}{|\mathbf{s}[i]|^2 |\mathbf{w}[i]|^2}] + 90^o term, \tag{4}$$

where w[k,i] are the detector data amplitudes - components of the vector **w**[i]. Both definitions of E_c are identical when $\mathbf{s}[i] = \mathbf{w}[i]$ With the new definition the $E_c[i]$ can be negative for noisy detector response. This is used to exclude such pixels from the reconstruction of the signal waveform.

The Ec distribution for background events depends on the regulator and it can be significantly different for the LH and LHV networks. For example, the LHV network effectively operates with different number of detectors depending on their relative sensitivity, sky location and frequency band. Despite the fact that formally this is a 3 detector network, the measurement can be performed with one detector ($E_c = 0$), 2 detectors (typically affected by hard regulator), 3 detectors (not affected by regulator) and all combinations in between. The overall EC destribution is a superposition of these different distributions that need to be normalized - the normalization coefficient is

$$R_c = \frac{1}{E_c} \sum_i E_c[i]/G_n[i]. \tag{5}$$

The R_c coefficient varies from 1 (hard regulator) and 0.5 (no regulator). For example, for the LH network $R_c = 1$.

3.1. Definition of the cWB detection statistic

The cWB detection statistic is defined as $\rho = \sqrt{ccE_c/N_{norm}/2}$, where cc is the network correlation coefficient. In the new version

$$\rho = \sqrt{R_c E_c/(2N_{norm})}, \xi^2 < 1, \qquad \rho = \sqrt{R_c E_c/(2\xi^2 N_{norm})}, \xi^2 > 1. \tag{6}$$

Similar to the CBC "new SNR" statistic it is penalized by the ξ^2, which is replacing cc in the previous definition of the statistic. This substitution is possible because of much more accurate estimation of ξ^2 in the new version.

3.2. Robust estimation of the network correlation coefficient

In cWB the network correlation coefficient is defined as $cc = E_c/(E_c + E_n)$, where E_n is the energy of the residual noise. Since $E_n \sim N_{DoF}$, the value of cc greatly depends on the time-frequency volume of the event. For this reason the cc threshold requires accurate tunning for different searches and networks. For example, the all-sky LH search has $cc > 0.8$., BBH searches - $cc > 0.7$, long duration search: $cc > 0.6$. In addition the LHV search will require different settings depending on the Virgo sensitivity. We have changed the definition of network correlation coefficient:

$$cc = E_c R_c/(E_c R_c + E_r + E_g - (K-1)N_{DoF}). \tag{7}$$

which produces about the same cc distribution for different signal morphologies and networks.

4. References

1. https://wiki.ligo.org/pub/Bursts/Cwb20170518/regulator.pdf

2. https://www.atlas.aei.uni-hannover.de/ waveburst/doc/cwb/man/The-WDM-packets.htmlThe-WDM-packets