LASER INTERFEROMETER GRAVITATION WAVE OBSERVATORY
-LIGO-
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Technical Report LIGO-T010125-00-D

A cross-correlation technique in wavelet domain for detection of stochastic gravitational waves

S.Klimenko, G.Mitselmakher

Distribution of this draft:

This is an internal working note
of the LIGO Project

LIGO Hanford Observatory
P.O. Box 1970;Mail Stop S9-02
Richland, WA 99352
Phone (509) 372-2325
Fax (509) 372-2178
E-mail: info@ligo.caltech.edu

LIGO Livingston Observatory
P.O. Box 1970;Mail Stop S9-02
Livingston, LA 70754
Phone (225)686-3100
Fax (225) 686-7189
E-mail: info@ligo.caltech.edu

California Institute of Technology
LIGO Project – MS 51-33
Pasadena, CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project – MS 20B-145
Cambridge, MA 01239
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

WWW.http://www.ligo.caltech.edu

Abstract

In this paper we describe a stochastic GW search technique in wavelet domain. It uses the rank and sign correlation tests, which allow calculate a significance of observed correlation for *non-Gaussian data*. As a part of optimal search technique we introduce a robust estimator for a GW detector noise spectral amplitude. It is insensitive to outliers in the data and allows to apply this search technique to *non-stationary data*. We also address the problem of *correlated noise* using a concept of correlation time scale.

1 Introduction

Recently there is an impressive progress in development of the gravitational wave (GW) interferometers [1-4]. One interferometer (TAMA, Japan) is collecting data and several more (LIGO, VIRGO and GEO) are about to start the data taking. In the frame of extensive scientific program, these interferometers will be used to search for the stochastic gravitational waves (SGW). The SGW might be produced by processes in the early universe and by large number of independent and unresolved GW sources [5-7]. It is exceptionally weak and a single detector cannot distinguish the SGW from an instrumental noise. However, if the SGW is correlated between several detectors, it can be detected using their cross-correlation. By integrating the cross-product of the detector output signals over a long period of time, we expect to enhance the signal to noise ratio (SNR) if the instrumental noise is not correlated.

This technique for detecting of the SGW using two or more gravitational wave detectors was described in [8,9]. It uses a linear correlation test [10], which allows estimate the significance of the observed correlation only if the instrumental noise is stationary and has a Gaussian distribution. In Section 2, we consider a robust correlation test that is free of these limitations. In Section 4, we illustrate how it works with two GW detectors.

A correlated instrumental noise arising from the environment could be a serious limitation for the cross-correlation technique described above. A weak background from seismic events, power supplies and other environmental sources may result in statistically significant correlation between the GW detectors, which may be misinterpreted as the SGW or affects the SGW upper limit. The SGW's contribution to cross-correlation depends on the relative orientation of the detectors. As suggested in [11], the contributions of the signal and the correlated background can be estimated separately by changing the orientation of one of the detectors, which is, in this case, the resonant bar detector ALLEGRO located in Louisiana State University (used in pair with the LIGO Livingston interferometer). Unfortunately, this method is not applicable to the GW interferometers, which are permanently located. In Section 3, we address the problem of correlated background for the GW interferometers using a concept of correlation time scale.

2 Robust correlation test

2.1 *Statement of the problem*

To characterize a correlation between two data sets x_i and y_i usually the linear correlation coefficient is used.

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2}\sqrt{\sum (y_i - \bar{y})^2}} \quad \textbf{(2.1)}$$

where $\bar{x}$ and $\bar{y}$ are means of the *x* and *y* distributions respectively. When the correlation is known to be significant, r is one conventional way to summarize its strength. However there is no universal way to compute r distribution in the case of null hypothesis, which is: *x and y are not correlated*. In other words *r* is a poor statistics to decide whether a correlation is statistically significant or whether one observed correlation is significantly stronger then another if the data is not Gaussian.

To solve this problem, often a rank statistic [10] is used (non-parametric Spearman's test). It has precisely known probability distribution function, which allows calculate the significance level of the observed correlation. The rank correlation test (RCT) is almost as efficient[1] as the linear correlation test (LCT) and potentially it is a good choice to be used for the SGW search. However, in this paper we do not discuss the rank correlation in details. The rank test is based on sorting algorithms, which are quite time consuming for large data sets. Instead we consider a robust correlation test (*sign test*), which is much simpler to use and easier to implement. Both correlation tests, rank and sign, can be used for the SGW search.

2.2 Description of the test

Let assume *x* to be produced by a random process with zero median. In cases when it's not true, a new statistics $x - \hat{x}$ can be calculated, where $\hat{x}$ is a median of the x data set. If to replace the value of each x_i with its sign u_j, the resulting first order statistics would be drown from a well known distribution function. It has zero mean and u_i take values +1 and –1. Given a second data set *y*, we repeat the procedure, replacing each value y_i by its sign v_i. Now we can build a statistics s_i, which is a product of u_i and v_i, for detecting a correlation between two sets. The correlation coefficient is simply a mean of the s_i statistics

$$s_{uv} = \overline{s_i} = \overline{u_i v_i}\,. \quad (2.2)$$

A value of s_{uv} near zero indicates that the variables *u* and *v* are uncorrelated. The significance of a nonzero value of s_{uv} is described by a binomial distribution.

$$P(n, r_{uv}) = \frac{n}{2} \cdot \frac{2^{-n} n!}{m!k!}, \quad m = \frac{n}{2}(1 - s_{uv}), \quad k = \frac{n}{2}(1 + s_{uv}), \quad (2.3)$$

where n is the number of s_i samples. This equation can be re-written using Sterling's approximation for factorials

$$P(n, s_{uv}) \approx \sqrt{\frac{n}{2\pi(1 - s_{uv}^2)}} \cdot (1 - s_{uv})^{-m}(1 + s_{uv})^{-k}\,. \quad (2.4)$$

For large *n,* the $P(n, s_{uv})$ can be also approximated by Gaussian distribution with variance *1/n*

$$P(n, s_{uv}) \approx \sqrt{\frac{n}{2\pi}} \cdot \exp\left(-\frac{n s_{uv}^2}{2}\right), \quad (2.5)$$

and a significance level is given by a complementary error function. Since, typically n is large, we will use this Gaussian approximation below in the text.

If we could know *a priori* the medians of x and y distributions, this test would be a *non-parametric* test. Since we estimate the medians from data, the test depends on the errors of $\hat{x}$ and $\hat{y}$, and s_{uv} may have a systematic shift from its true mean. It can be shown that this dependence is very weak and, hence, the test is robust. Indeed, the mean of u (and v) distribution has a variance $1/n$. Respectively, the mean of s_i distribution experiences systematic fluctuations with variance $2/n^2$, which is much less then the intrinsic variance of s_{uv} given by the test ($1/n$). We could say that for large n the *sign correlation test (SCT)*, is *almost non-parametric*.

2.3 Comparison of correlation tests

The SCT has been studied using Gaussian signals and noise with various distributions. The goal of the study was to estimate the SCT efficiency compare to the linear correlation test (LCT), which is supposed to be one of the most efficient correlation tests. For comparison we also estimated the efficiency of the rank correlation test (RCT).

For the data sets, first we generated two time series n_x and n_y with uncorrelated white noise. We tried various noise probability distribution functions: Gaussian, Gaussian with tails, asymmetric Gaussian and uniform (see Figure 1). Then a Gaussian signal g was added to both time series, so the data (x and y) is a sum of uncorrelated noise and correlated signal

$$x = n_x + g, \quad y = n_y + g. \quad (2.6)$$

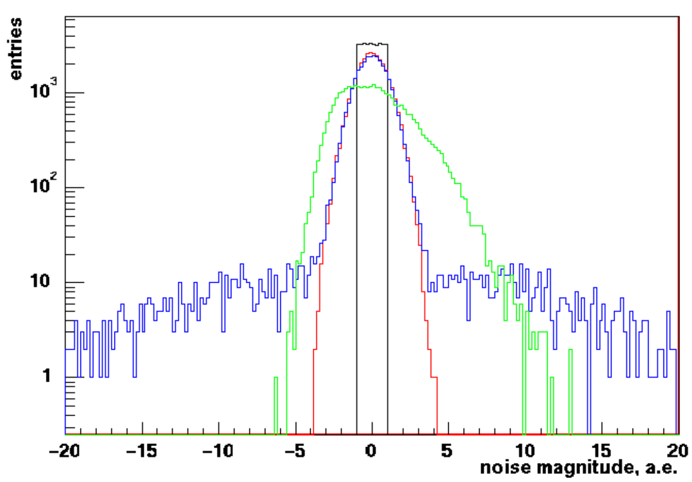

Figure 1. Distribution of noise used to compare different correlation tests: Gaussian (red), Gaussian with tails (black), asymmetric (blue), uniform (green).

The amplitude of the signal was relatively small (signal to noise ratio (SNR) <0.5), where the SNR was calculated as a ratio of standard deviations of the signal and noise distributions.

To compare the correlation tests, the correlation coefficients (R) were calculated. Figure 1 shows the correlation coefficients as a function of SNR for Gaussian signals and different correlation tests (LCT, RCT, SCT).

---

[1] See discussion of the correlation test efficiency in section 2.3

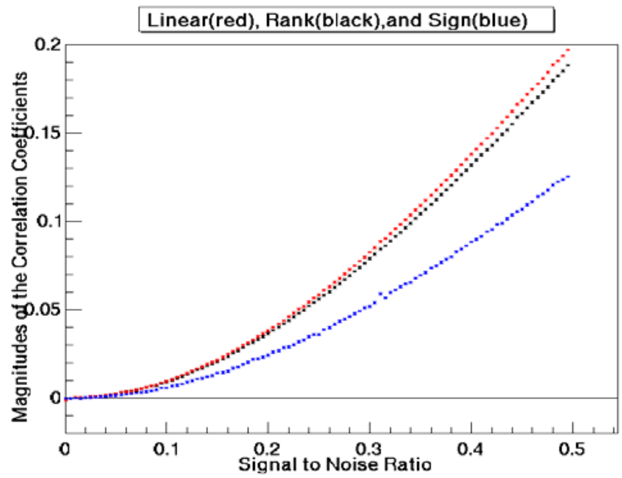

Figure 2. Correlation coefficients as a function of SNR for Gaussian noise: LCT(red), RCT (black), SCT (blue).

To estimate the SCT efficiency we calculate the ratio $\overline{R}_{SCT} / \overline{R}_{LCT}$ for different types of noise. This ratio (or the sign correlation efficiency) doesn't depend much on the signal to noise ratio and for the Gaussian noise it is around 64%. For comparison, similar ratio for the RCT is 95%. Figure 3 shows the sign correlation efficiency for different types of noise.

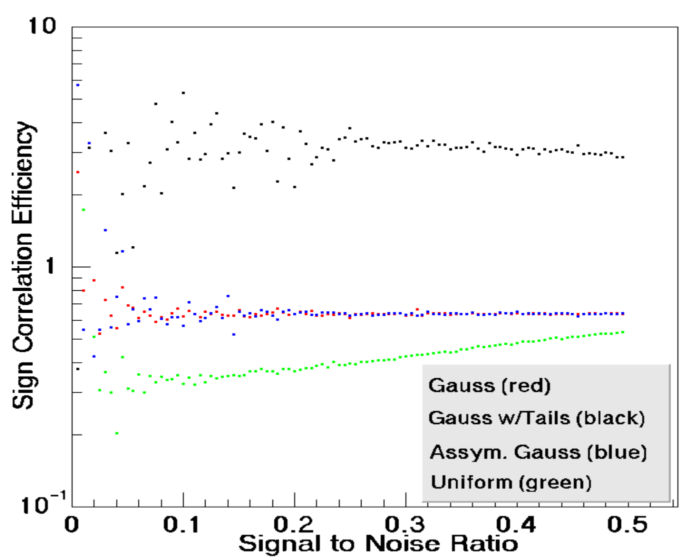

Figure 3. The sign correlation efficiency as a function of SNR: Gaussian (red), Gaussian with tails (black), asymmetric (blue), uniform (green).

Typically, if applied to the same data, the SCT will detect correlation with less significance level then the LCT. For Gaussian noise, to achieve the same significance level with the SCT we need to use about 2.4 times more data (1.1 for the RCT), which means effective loss of data compare to the LCT. However, for Gaussian noise with tails, which is the most typical type of noise, the SCT shows better efficiency then the LCT.

3 Correlated noise

3.1 Statement of the problem

In case of two GW detectors, let say *L* and *H*, the output of each detector is a mixture of the SGW signal (h) and noise (n)

$$s_{L,H} = h_{L,H} + n_{L,H} \qquad (3.1)$$

Assuming no correlation between signal and noise, the cross-correlation of the data s_L and s_H is

$$r = \overline{h_L h_H} + \overline{n_L n_H}\,. \qquad (3.2)$$

Due to correlated noise, the noise cross-correlation term may not be zero and the correlation test may find correlation, if even no signal is present in the data. For example, if detector output is dominated by power supply harmonics, which are very stable in time, the correlation test will always find a significant correlation between the detectors. Combined with correlated signal the noise may result in a wrong upper limit on the SGW. The correlation test may also fail to find a correlation if the noise is anti-correlated.

The problem with the correlation tests described above is, that applying a test we check if <u>*any*</u> correlation is present in the data. However, different correlation processes may be distinguished by their *correlation time scale* and the correlation tests can be modified accordingly. Both, rank and sign correlation tests can be modified to accommodate the correlated noise. Below we show it on example of the sign correlation test.

3.2 Correlation time scale

In general, there could be a correlation between samples of the sign statistics s_i and $s_i(t)$ is a step function of time. It is fully described by its normalized autocorrelation function $a(\tau) = p_+(\tau) - p_-(\tau)$, where p_+ (p_-) is the probability to find the same (opposite) signs of s_i and s_j separated by time τ. For $a(\tau)$ it holds that $a(0) = 1$ and $a(\tau > T_s) = 0$[2], where T_s is a *correlation time scale (CTS)*. Namely, any samples s_i and s_j, separated by time τ greater then T_s, are not correlated.

The function $a(\tau)$ depends on a correlation process between the data sets. The time scale T_s characterizes the function's width. For example, if the statistics s_i is a white noise, then $T_s = \Delta t$, where Δt is the sampling interval. Different correlation processes due to instrumental noise may have different correlation time scales, which may be greater then Δt. For environmental noise we expect a correlation time scale to be finite, thought the SGW correlation time scale is assumed to be infinite. In other words, if a data set is long enough, such a time T_s can be selected, that only a faint SGW signal will contribute to the function $a(\tau > T_s)$[3].

The sign correlation test (as well as the linear and rank tests) tells us if there is a correlation with any time scale equal or larger then Δt. It does not use any noise model[4] and the null hypothesis reads accordingly – *data sets are not correlated*. So, if a correlated noise is

[2] We assume no noise process exists with an infinite correlation time scale.
[3] See Section 5.3 for more detailed discussion of correlated noise.
[4] So far we didn't use any signal model as well.

present in the data, a different null hypothesis should be used:

H_0: Data sets are not correlated at time scale greater then some scale T_s and the observed correlation coefficient is described by noise with the time scale less then T_s.

Respectively, a correlation test should be modified to incorporate this new parameter.

To modify the test, lets divide the correlation statistics s_i on subsets of m samples long (time duration $T_m = m \cdot \Delta t$) and calculate the correlation coefficients s_{uv} separately for each subset. Lets also calculate the product of T_m and the variance of the correlation coefficients

$$\Delta T(T_m) = T_m \cdot \overline{s_{uv}(T_m)^2}\,, \qquad (3.3)$$

which is the *effective sampling interval*. The meaning of this name is clear from the following.

If some correlated noise with time scale T_s is present in the data and we increase T_m, the function $\Delta T(T_m)$ will increase first and then saturates Indeed if $T_m > T_s$, samples in the beginning and the end of a subset are not correlated. Then, in the Gaussian approximation,

$$\overline{s_{uv}(T_m)^2} = const \cdot \frac{\Delta t}{T_m}, \qquad (3.4)$$

and the mean of the ΔT distribution is independent on T_m

$$\overline{\Delta T}(T_m > T_s) = const \cdot \Delta t\,. \qquad (3.5)$$

For example, Figure 4 shows $\Delta T(T_m)$ for the simulated white noise and for the cross-correlation of the LIGO Livingston and Hanford power supply monitors (60Hz). For white noise it holds that $\overline{\Delta T}(t_m \geq \Delta t) = \Delta t$ ($\Delta t = 1/64\,\mathrm{sec}$), which can be clearly seen from the figure. For power monitors the $\Delta T(T_m)$ saturates at $T_m > 60\,\mathrm{sec}$, which means that at this time scale the sign correlation statistics is effectively a white noise with a sampling interval of $\overline{\Delta T} \approx 10\,\mathrm{sec}$.

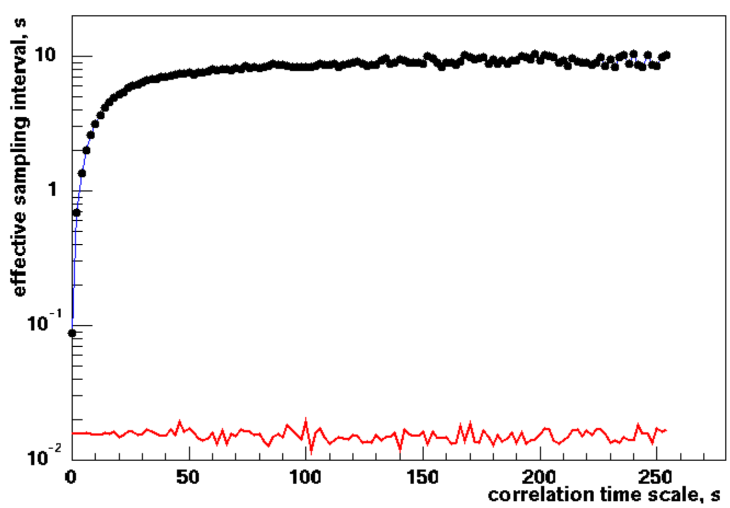

Figure 4. Effective sampling interval for the LHO/LLO power monitor sign cross-correlation (black) and simulated white noise (red).

When $\Delta T(T_s) < T$, where T is the time duration of the data, the correlation coefficients s_{uv}

still have a binomial distribution

$$P(\widetilde{n}, s_{uv}) = \frac{\widetilde{n}}{2} \cdot \frac{2^{-\widetilde{n}} \widetilde{n}!}{m!k!}, \quad m = \frac{\widetilde{n}}{2}(1 - s_{uv}), \quad k = \frac{\widetilde{n}}{2}(1 + s_{uv}), \qquad (3.6)$$

where $\widetilde{n}$ is the *effective number of samples*

$$\widetilde{n} = n \frac{\Delta t}{\Delta T(T_s)}. \qquad (3.7)$$

In the Gaussian approximation, the variance of s_{uv} distribution is $1/\widetilde{n}$. One can see that correlated noise effectively increases variance, resulting in a lower significance level

$$SL = erf\left(\frac{n}{2} \frac{\Delta t}{\Delta T(T_s)} s_{uv}^2\right). \qquad (3.8)$$

It is convenient also to introduce a *reduced correlation coefficient* $\widetilde{s}_{uv}$, which is distributed with variance $1/n$

$$\widetilde{s}_{uv} = s_{uv} \sqrt{\frac{\Delta t}{\Delta T(T_s)}}. \qquad (3.9)$$

Usually, we don't know a time scale of the noise *a priori*. In principle, we may select any time scale T_s between Δt and T [5]. Varying T_s, we can calculate the effective sampling interval and thus the significance level. If starting with some T_s, the significance level (3.9) is greater then, let say, *5%*, we choose the hypothesis H_0 otherwise we choose the alternative hypothesis:

H_1 : *Data sets are correlated at time scale greater then scale* T_s.

In this case, we can calculate a false alarm rate (H_0 is true, but we choose H_1), but we are not able to calculate a false dismissal rate (H_1 is true, but we choose H_0), because so far a model for signal was not specified.

4 Correlation in wavelet domain

4.1 Wavelet transforms

The term wavelet is usually associated with a function $\psi \in L_2(R)$ [6] such that the translation and dyadic dilation of $\psi \in L_2(R)$ constitute an orthonormal basis of $L_2(R)$ [12,13]. For discrete wavelet transform the basis is

$$\psi_{mn}(t) = 2^{n/2} \psi(2^n t - m), \; n, m \in Z, \qquad (4.1)$$

where ψ satisfies

$$\int \psi(t) dt = 0. \qquad (4.2)$$

The important property of wavelets is a time-frequency localization of their basis. It allows the time-frequency representation of data in wavelet domain, similar to windowed Fourier transform. The result of a wavelet transform of data x is an array of wavelet coefficients p_{mn},

[5] Actually, we have to limit the time scale T_s between Δt and T_L, where $T_L << T$.

[6] Space of all square-integrable functions.

where m is a time index and n is a scale (or layer) index. Applied to wavelet data, the correlation can be estimated as a function of the layer index, which represents different frequency bands of the data x. Each layer can be considered as a time series with the sampling interval $\Delta t_n = 2^n \Delta t$.

The same as for the original data x, we don't know the exact probability distribution function for coefficients p as well. However, if consider the detail wavelet coefficients, usually their distribution is "more Gaussian" (see, for example, Figure 5).

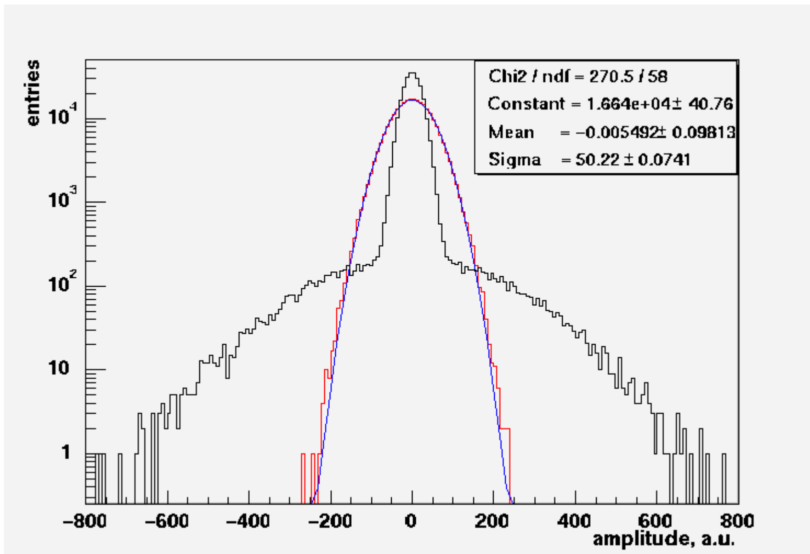

Figure 5. Distribution of white noise with tails in time (black) and wavelet (DAUB4) domains (red). Blue curve is a Gaussian fit of data in wavelet domain.

Due to condition (Eq.4.2) the distribution of wavelet data usually has zero mean and it is also symmetric, which makes convenient to perform the sign transform in wavelet domain.

To apply the wavelet method to the SGW search, we modify the linear correlation test described in [14]. First (Section 4.2), we calculate the linear cross-correlation in wavelet domain and then (Sections 4.3, 4.4) we consider the sign cross-correlation.

4.2 Cross-correlation in wavelet domain

The cross-correlation between two detectors is

$$S = \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt' s_L(t') s_H(t) Q(t - t', \Omega_L, \Omega_H), \qquad (4.3)$$

where T is the integration time and Ω_H (Ω_L) is the orientation of H (L) interferometer. The integration kernel Q is selected to maximize the correlation due to the stochastic GW signal. Decomposing s_L and s_H in wavelet domain with a discrete wavelet transform

$$s_L(t) = \sum_{k,l} p_{kl} \psi_{kl}(t), \qquad s_H(t) = \sum_{n,m} q_{mn} \psi_{mn}(t) \qquad (4.4)$$

where ψ_{ij} is the orthonormal basis of wavelet functions, we can rewrite 4.3

$$S = \sum_{nm} \sum_{k,l} p_{kl} q_{mn} I_{kl,mn}, \qquad (4.5)$$

$$I_{kl,mn} = \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt' \psi_{kl}(t')\psi_{mn}(t) Q(t-t', \Omega_L, \Omega_H). \qquad (4.6)$$

We can also write double integral $l_{kl,mn}$ in Fourier domain. Using the time localization of wavelet functions it follows that

$$I_{kl,mn} = \int_{-\infty}^{\infty} df \hat{\psi}_{kl}(f)\hat{\psi}^*_{mn}(f) Q(f, \Omega_L, \Omega_H) \qquad (4.7)$$

where $\hat{\psi}(f)$ and $Q(f, \Omega_L, \Omega_H)$ are the Fourier transforms of $\psi(t)$ and $Q(t, \Omega_L, \Omega_H)$. Note, if $Q(t-t') = \delta(t-t')$, due to the orthogonality of wavelet basis, it holds that $I_{kl,mn} = \delta_{km}\delta_{\ln}$, where δ_{km} ($\delta_{\ln}$) is Kronecker delta. It is not true for an arbitrary kernel, but because of good localization of wavelet functions in time and frequency the most of the $I_{kl,mn}$ terms are equal zero. Because of the frequency localization, we may neglect terms with $l \neq n$ and rewrite the eq. as follows

$$S \approx \sum_n S_n, \quad S_n = \sum_{k,m} p_{kn} q_{mn} I_{kn,mn} \qquad (4.8)$$

where S_n is a cross-correlation for the n^{th} layer. Since wavelet functions for each layer are produced by translation in time of the same mother function ψ_n, the Fourier integral for $I_{kn,mn}$ is

$$I_{kn,mn} = I_n(\tau) = \int_{-\infty}^{\infty} df \hat{\psi}_n(f)\hat{\psi}^*_n(f) Q(f, \Omega_L, \Omega_H) \exp(-j2\pi f\tau)., \qquad (4.9)$$

where τ is a time shift between coefficients p and q ($\tau = \Delta t_n(k-m)$). Assuming that the noise of each detector is much larger in magnitude then the SGW signal, the noise mean square for the n^{th} layer is

$$\sigma^2_{nL} = \frac{1}{N_n}\sum_{m=1}^{N_n} p^2_{mn}, \qquad \sigma^2_{nH} = \frac{1}{N_n}\sum_{m=1}^{N_n} q^2_{mn}. \qquad (4.10)$$

where N_n is the number of samples in the layer. Introducing the coefficient of linear correlation for layer n

$$r_n(\tau) = \frac{1}{N_n \sigma_{nH}\sigma_{nL}} \sum_{m=1}^{N_n} p_n(m\Delta t_n) q_n(m\Delta t_n - \tau), \qquad (4.11)$$

the cross-correlation sum S calculated in wavelet domain is

$$S = \sum_n N_n \sigma_{nL}\sigma_{nH} \sum_\tau I_n(\tau) \cdot r_n(\tau) = \sum_{n,\tau} w_n(\tau) r_n(\tau). \qquad (4.12)$$

So far the cross-correlation S is equivalent to the result obtained in the Fourier domain [14]. Similarly we can use the optimal kernel to calculate $I_n(\tau)$ and hence $w_n(\tau)$

$$Q(f, \Omega_L, \Omega_H) = \frac{|f|^{-3}\, \Omega_{GW}(f)\gamma(f, \Omega_L, \Omega_H)}{P_L(f) P_H(f)}, \qquad (4.13)$$

where $\Omega_{GW}(f)/f$ is proportional to the GW differential energy density, γ is the detector overlap reduction function [15] and P_L, P_H are spectral densities of the detector noise. Then the full expression for the $w_n(\tau)$ is

$$w_n(\tau) = N_n \int_{-\infty}^{\infty} df \left|\psi_n(f)\right|^2 \left|f\right|^{-3} \Omega_{GW}(f)\gamma(f,\Omega_L,\Omega_H)\frac{\sigma_{nL}}{P_L(f)}\cdot\frac{\sigma_{nH}}{P_H(f)}\exp(-j2\pi f\tau). \quad (4.14)$$

The equation shows that the cross-correlation S is a weighted sum of linear correlation coefficients $r_n(\tau)$. For a pair of GW detectors, the optimal weight coefficients can be calculated for a selected SGW model ($\Omega_{GW}(f)$) and using an appropriate noise spectral density estimator. However, as we have mentioned above two questions remain:

a) What is the actual distribution of the correlation coefficients $r_n(\tau)$?
b) How to calculate an optimal weight coefficients if the detector noise is non-stationary or correlated?

To solve those problems, we suggest to use a robust correlation test, like RCT or SCT. In the next section we show how it works for the sign correlation test. The rank correlation test can be implemented in a similar way. The problem of correlated noise could be partially solved, by taking in to account the correlation time scale (see Section 4.4).

4.3 Uncorrelated noise

To apply the sign test, we simply replace the coefficients $r_n(\tau)$ in the cross-correlation sum S (Eq.4.12) with the corresponding sign (or rank) correlation coefficients $s_n(\tau)$. It may reduce the efficiency of the test, but we gain confidence in calculation of the significance level. To keep the weight coefficients optimal, the sign correlation efficiency (see Section 3.1) should be taken into account. The sign correlation efficiency (ε_n) may be different for different wavelet layers. Then, the optimal weight coefficients are

$$\tilde{w}_n(\tau) = w_n(\tau)\varepsilon_n \ , \quad (4.15)$$

and the sign cross-correlation is

$$S_s = \sum_{n,\tau} \tilde{w}_n(\tau) s_n(\tau). \quad (4.16)$$

As it was mentioned in section 2.2, the correlation coefficients $s_n(\tau)$ are normally distributed with variance *1/N*$_n$. Then, the variance of S_s is

$$\mathrm{var}(S_s) = \sum_{n,\tau} \frac{1}{N_n} \tilde{w}_n^2(\tau). \quad (4.17)$$

In many cases we can consider that all efficiencies ε_n are the same (around 65%) and just reduce the signal to noise ratio

$$SNR^2 = \frac{S_s^2}{\mathrm{var}(S_s)}. \quad (4.18)$$

However if the noise distribution has tails, the efficiencies ε_n can be very different (see Figure 2).

We can include the ε_n in the expression for $w_n(\tau)$ (Eq.4.14) and introduce a *robust noise spectral amplitude*

$$A_I(f) = \frac{P_I(f)}{\sigma_{nI}\sqrt{\varepsilon_n}}, \quad I{=}L,H\,. \quad (4.19)$$

Then the optimal weight coefficients $\tilde{w}_n(\tau)$ are

$$\widetilde{w}_n(\tau) = N_n \int_{-\infty}^{\infty} df \left|\psi_n(f)\right|^2 |f|^{-3} \frac{\Omega_{GW}(f)\cdot\gamma(f,\Omega_L,\Omega_H)}{A_L(f)\cdot A_H(f)} \exp(-j2\pi f\tau). \quad (4.20)$$

One can see that only the amplitudes A_L and A_H should be estimated from the experimental data to calculate the optimal weight coefficients.

The robustness of amplitudes A_I can be shown on example of Gaussian noise with tails. For simplicity we use a white noise, then $P_I(f) = const$. If, let say, we have a Gaussian noise with fixed variance σ_g^2 and vary contribution from the tails, the value of spectral density P is proportional to the total noise variance σ_n^2 and it increases, if we increase the tail contribution (see Table 1).

σ_n/σ_g	P	A
1.0	0.45	0.0266
1.45	0.94	0.0274
2.31	2.40	0.0273

Table 1. Dependence of the noise SD and noise RSD on tail contribution

At the same time one can see that the amplitude A remains constant. It makes calculation of the optimal weight coefficients much more robust if the tails are non-stationary.

To find the efficiency ε_n we use the fact that it is constant as a function of the SNR for all reasonable distribution functions of the noise (see Figures 2,3). Since we assume the SGW signal to be Gaussian, we could estimate the efficiencies s_n by adding a Gaussian signal g_m to the data p_{mn} and q_{mn} in the wavelet domain. Calculating the LCT and SCT correlation coefficients, we can find the efficiencies s_n for relatively large SNR, when the signal g_m dominates, and then extrapolate it to small SNR, using the assumption $\varepsilon(SNR) = const$.

4.4 Correlated noise

Compare to the Section 3.1, now we are dealing with a set of correlation coefficients $s_n(\tau)$. Each coefficient $s_n(\tau)$ is a mean of the corresponding sign correlation statistics $s_i(n,\tau)$. If some correlated noise is present in the data, the different statistics $s_i(n,\tau)$ may not be statistically independent. Similar to the sign statistics autocorrelation function (see Section 3.1) we may introduce a cross-correlation function for two statistics $s_i(n_1,\tau_1)$ and $s_i(n_1,\tau_1)$. If samples of $s_i(n,\tau)$ are not correlated at the time scale $T_s(n,\tau)$, then samples of any two statistics $s_i(n_1,\tau_1)$ and $s_i(n_2,\tau_2)$, separated by time $\max(T_s(n_1,\tau_1),T_s(n_2,\tau_2))$ do not correlate as well. It means, that at this time scale the correlation coefficients are statistically independent and they have a binomial distribution. In the Gaussian approximation the optimal cross-correlation sum is

$$\widetilde{S}_s = \sum \widetilde{w}_n(\tau)\sqrt{\frac{\Delta t}{\Delta T(n,\tau)}}s_n(\tau) = \sum_{n,\tau} \widetilde{w}_n(\tau)\widetilde{s}_n(\tau), \qquad (4.21)$$

where $\widetilde{s}_n(\tau)$ are the reduced correlation coefficients (see Section 4.1) and $\Delta T(n,\tau)$ are the effective sampling intervals. The variance of $\widetilde{S}_s$ is equal to the variance of S_s (see Eq.4.17) and the SNR is

$$SNR^2 = \left(\sum_{n,\tau} \widetilde{w}_n(\tau) \widetilde{s}_n(\tau) \right)^2 \Bigg/ \sum_{n,\tau} \frac{1}{N_n} \widetilde{w}_n^2(\tau). \qquad (4.22)$$

One can see that the correlated noise effectively reduces the signal to noise ratio.

This approach for correlated noise works if the noise time scale is much less then the SGW time scale[7] and it fails if the noise time scale is comparable to the SGW time scale. Of course, the first type of noise does some harm – it reduces the significance of correlation, but the second type of noise can make the detection of the SGW impossible.

5 Conclusion

The SGW signal should manifest itself in the cross-correlation of several detectors. In this paper we discussed the problem (i) of estimating the significance of the cross-correlation if the detector noise is non-Gaussian and the problem (ii) of correlated noise.

Different correlation techniques can be used to calculate the detector cross-correlation. We discussed the optimal signal processing in wavelet domain using the rank and sign correlation tests, which allow calculate the significance of the cross-correlation if the detector noise is non-Gaussian. The rank correlation is a non-parametric test and it is almost as efficient as the linear correlation test. It is based on sorting algorithms and could be quite time consuming for large data sets. The sign correlation test is robust and simple to use, however it is less efficient (65%) compare to the linear and rank correlation tests. Using the sign correlation test we also introduced a spectral density estimator, which allows a robust calculation of the noise spectral density if the detector noise distribution has outliers.

We addressed the problem of correlated noise using the concept of correlated time scale. Namely, different sources of correlated noise can be distinguished by their correlation time scale. In this paper we developed a method, which allows estimate the significance of the correlation in the presence of correlated noise. The concept of correlation time scale is not an ultimate solution of the problem, however it allows a consistent calculation of the cross-correlation if the correlated noise time scale is much less then the SGW time scale.

This approach fails to work for noise sources with a very long time scale, but most likely there should be very few sources like that. For example, it could be a noise due to magnetic field fluctuations [15], the power system synchronization, seismic events coming from locations equally distant from the detectors, and also the data sampling and data processing artifacts. In this case we need to study this sources (for example using the environmental data channels) and build a consistent noise model to estimate correctly the correlated noise contribution.

Acknowledgments

We are grateful to Bruce Allen for suggestions, which considerably improved the paper and Bernard Whiting for useful comments and suggestions regarding the text. We also thank our REU student Jason Castiglione who helped with calculations involving the rank and sign correlation tests. This work was supported by NSF grant PHY-0070854.

[7] We assume that we are not limited by the integration time.

References

1. A.Abramovici et al., Science 256, 325 (1992).
2. B. Caron et al., in Gravitational Wave Experiments, proceedings of the Edoardo Amaldi Conference, World Scienti_c, 86 (1995).
3. K. Danzmann et al., in Gravitational Wave Experiments, proceedings of the Edoardo Amaldi Conference, World Scienti_c, 100 (1995).
4. K. Tsubono, in Gravitational Wave Experiments, proceedings of the Edoardo Amaldi Conference, World Scienti_c, 112 (1995).
5. P. F. Michelson, Mon. Not. Roy. Astron. Soc. 227, 933 (1987).
6. N. Christensen, Phys. Rev. D46, 5250 (1992).
7. E. Flanagan, Phys. Rev. D48, 2389 (1993).
8. B.Allen, gr- qc/ 9604033 v3 30 Sep 96
9. B.Allen and J.D.Romano, gr- qc/ 9710117 v1 27 Oct 1997
10. W.H.Pres et al, Numerical Recipes, Cambridge University press, 1992
11. A.Lazzarini and L.S.Finn, PACS numbers: 04.80.Nn, 04.30.Db, 11.25.-w, 98.80.Cq, 98.80.Es
12. I. Daubechies, Ten lectures on wavelets. Philadelphia, PA, 1992.
13. B. Vidakovic, Statistical modeling by wavelets. John Wiley & Sons, Inc, 1999
14. J.Romano, Correlation test based on truncating the tails in data distribution. (private communication)
15. N. Christensen, Phys. Rev. D46, 5250 (1992).

$$\Omega_{GW}(f) = \frac{1}{\rho_c}\frac{d\rho_{GW}(f)}{d\ln(f)} \qquad \rho_c = \frac{3c^2 H_0^2}{8\pi G} \approx 1.6 \cdot 10^{-8}\,\frac{ergs}{cm^3}\; H_0 = 100\,\frac{km}{\sec Mpc}$$