x-correlation in wavelet domain for detection of stochastic gravitational waves

arXiv: gr-qc/0208007v1 Aug 2, 2002

S.Klimenko, University of Florida

Outline

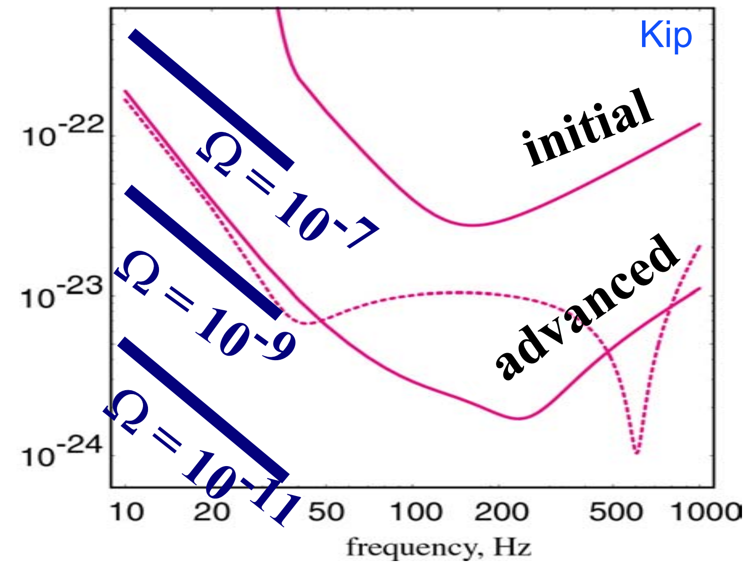
- Introduction
- Optimal cross-correlation
- Correlation tests
- Correlated noise
- robust x-correlation in wavelet domain
- Conclusion

Introduction

- **Stochastic Gravitational Waves**
  - **from early universe or/and large number of unresolved sources**

(GW energy density)/(closure density)

$$\Omega_{GW} < 10^{-5}$$



- **Detection of SGW**
  - **x-correlation of detector output signals** $s_L(t)$ *and* $s_H(t)$

$$S = \int_0^T dt \int_0^T dt' s_L(t') s_H(t) Q(t - t', \Omega_L, \Omega_H)$$

  - T **– observation time,** Q **- optimal kernel**
  - Ω_H (Ω_L) ***is the orientation of H (L) interferometer***

Optimal Cross-Correlation

- **x-correlation in Fourier domain**

B. Allen and J.D. Romano,
Phys. Rev. D59, pp. 102001-102041, 1999

$$S = \int_{-\infty}^{\infty} df\, \tilde{s}_H(f)\tilde{s}_L^*(f) Q(f,\Omega_L,\Omega_H)$$

  - **Optimal kernel:**

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

    - ✓ Ω_{GW} – **SGW strength**
    - ✓ P_L, P_H **- spectral densities of detector noise**
    - ✓ γ **– detector overlap function** (E. Flanagan, Phys. Rev. D48, 2389 (1993))

- **Questions:**
  - **What is the *distribution of S* if noise is not Gaussian?**
    Allen, Creighton, Flanagan, Romano, PRD 65, 122002, 2002
  - **What if *S* is affected by correlated noise?** (LLO-Allegro)
  - **How Q is affected?**

Correlation Tests

- **linear**

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

use knowledge of r distribution to perform the test

  - parametric: no universal way to compute ***r*** distribution
  - if data is not Gaussian, ***r*** is a poor statistics to decide
    - ✓ **correlation is statistically significant**
    - ✓ **one observed correlation is stronger then another.**

- **rank**
  - non-parametric: exactly known ***r*** distribution
  - CPU inefficient for large data sets

- **sign**
  - primitive version of the rank test

Sign Correlation Test

- Sign transform: $u_i = sign(x_i - \hat{x})$ ± 1
  - $\hat{x}$ - **sample median of *x***
- Sign corr. statistic: $s_i = sign(x_i - \hat{x}) \cdot sign(y_i - \hat{y})$
- Correlation coefficient ρ: $\rho = mean(s_i)$
- Distribution of ρ (n – number of samples):
  - **Gaussian (large n):**
  - **variance = *1/n***

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

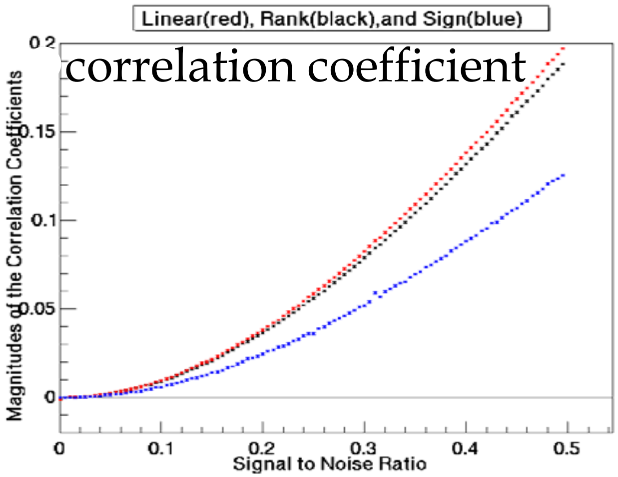
- very robust:
  - error from $\hat{x}$ and $\hat{y}$ $\sim 2/n^2$, much less then ***var(ρ)=1/n*** for large *n*
  - sensitivity? correlation between s_i samples?

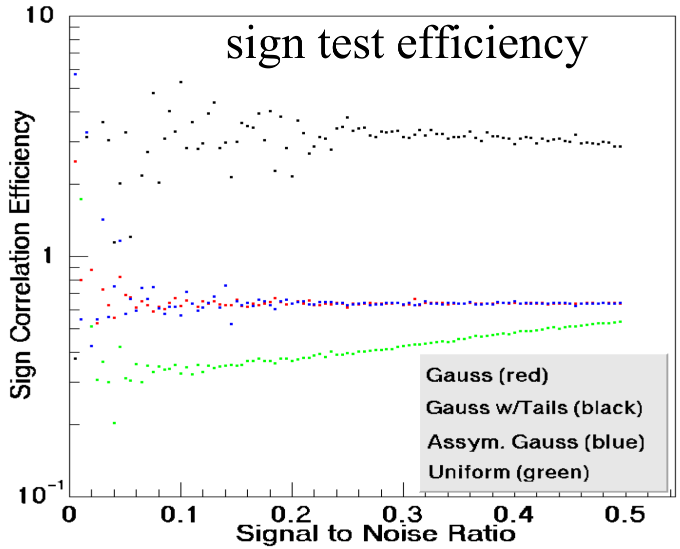
Comparison of Correlation Tests

- **Data:** simulated uncorrelated noise (n) + Gaussian signal (g)

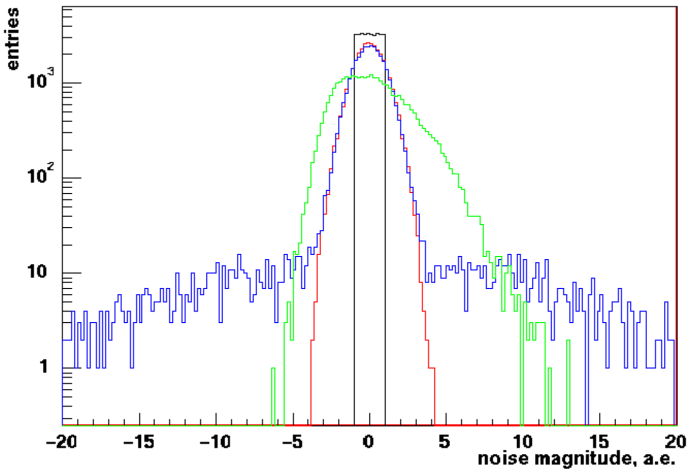
$$x = n_x + g, \qquad y = n_y + g$$

- **Test efficiency:** $\varepsilon_s = r_s / r_L$
  - for Gaussian noise
    - ✓ rank test efficiency – 95%
    - ✓ sign test efficiency - 64% (2.5 loss of data)
  - independent on SNR



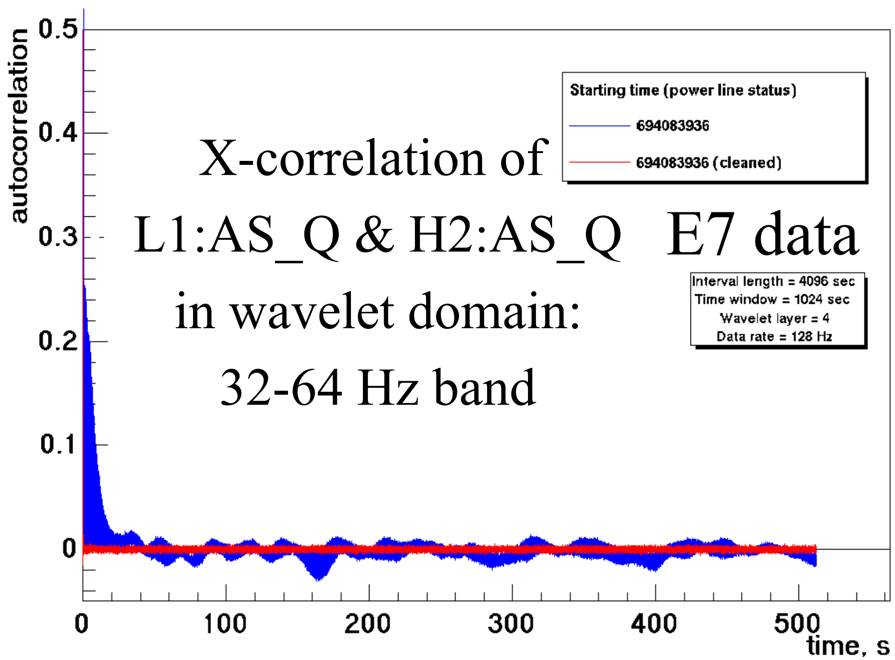


noise distributions



Correlated Noise

- **What does it mean that data is affected by correlated noise?**
  - **ρ mean value is biased:** $$\langle \rho \rangle = \langle h_L + n_L, h_H + n_H \rangle = \langle h_L, h_H \rangle + \langle n_L, n_H \rangle$$
  - **ρ variance is affected:** $$\text{var}(\rho) = \langle \rho^2 \rangle - \langle \rho \rangle^2$$
- **Samples of sign correlation data *{s_i}* can be correlated**
  - ***{s_i}* is generated by some random process *$s(t)$* with autocorrelation function *$a(\tau)$***
- ***$a(\tau)$***
  - **takes into account second order statistic *$P(s_i,s_j)$***
  - **a measure of correlated noise.**



S.Klimenko, LSC August 2002

Noise Model

- **uncorrelated noise**
  - **autocorrelation function:** $a(0)=1,\ a(\tau \geq \Delta t)=0$
  - **null hypothesis:** ***data sets are not correlated***
  - **variance:** $\mathrm{var}_0(\rho) = \frac{1}{n}$
- **correlated process with time scale <*T_s***
  - **autocorrelation function:** $a(\tau < T_s) = a_n(\tau),\ a(\tau > T_s) = 0$
  - **null hypothesis:** ***data sets are not correlated at time scale >T_s***
  - **variance:** $\mathrm{var}_{Ts}(\rho) = \frac{1}{n}\nu,\quad \nu = 1 + \frac{2}{n}\sum_{m=1}^{T_s/\Delta t}(n-m)a_n(m\Delta t)$
- **calculation of *var(ρ)*, depending on the noise model.**

ν

- **variance ratio**

$$\nu(T_s) = \frac{\mathrm{var}_{Ts}(\rho)}{\mathrm{var}_0(\rho)}$$

  - **ν is a measure of correlated noise, or quality of data.**
  - **ν times more data needed to reach same CL as for uncorrelated noise.**
  - **If ν is too large, the noise should be removed, if possible**

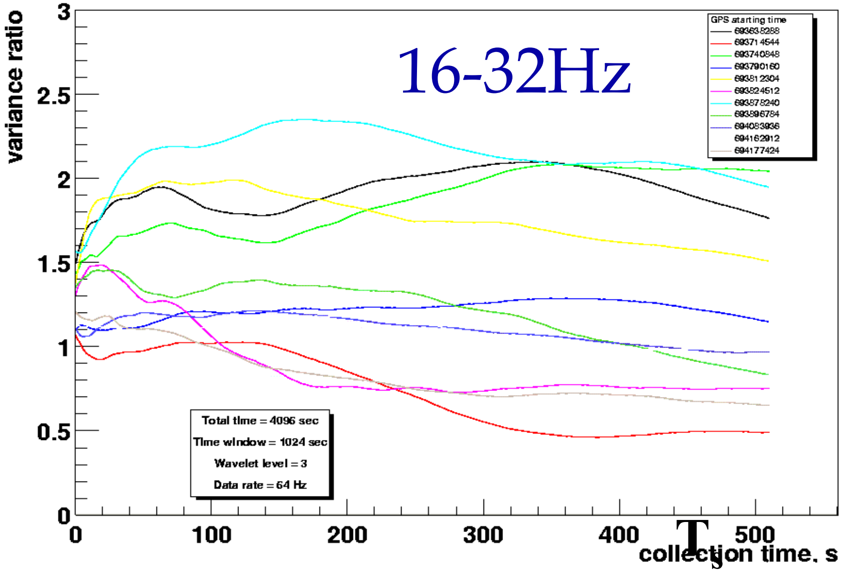
- **correlation time**

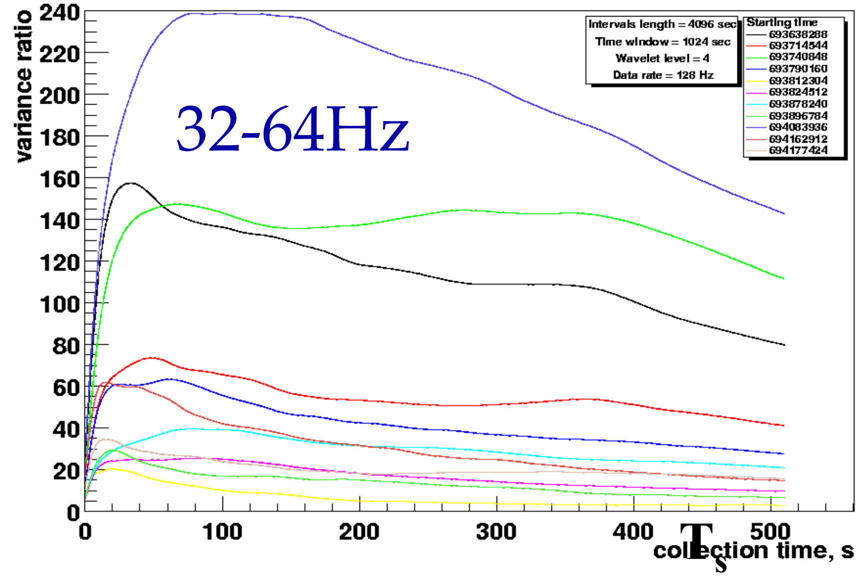
$$T_c = \int_0^\infty a(t)dt \approx \frac{\nu\Delta t}{2}$$

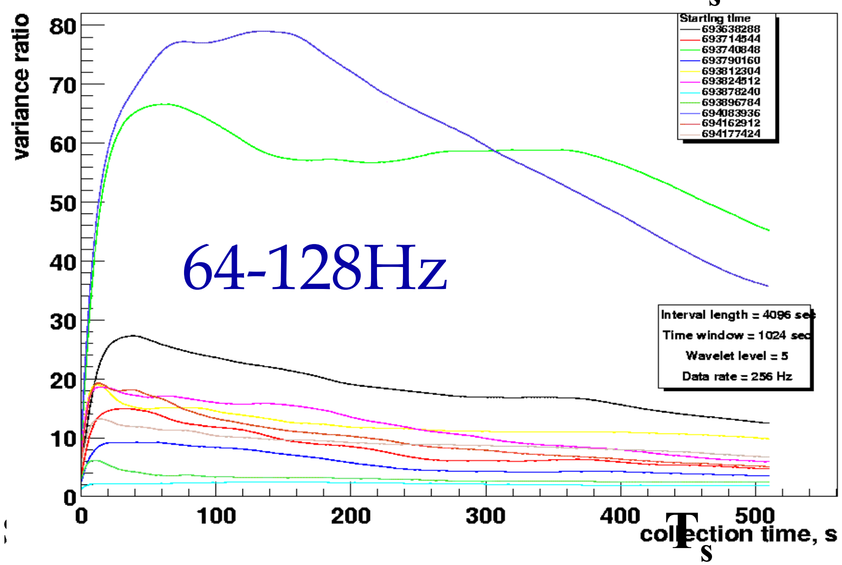
$$\mathrm{var}(\rho) = \frac{1}{n}\nu \approx \frac{2T_c}{T},$$

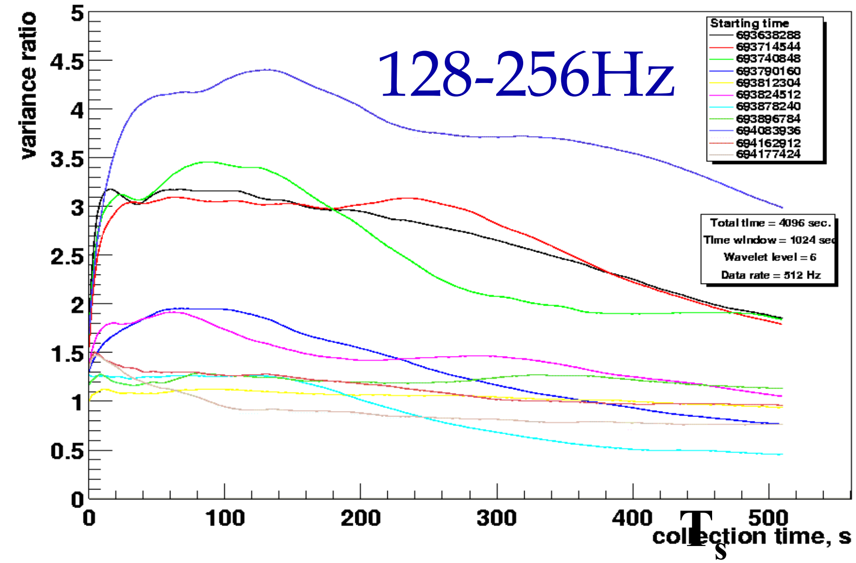
Variance Ratio (T_s)

- **L1xH2: 11 data segments 4096 sec each (total 12.5 h of E7 data)**



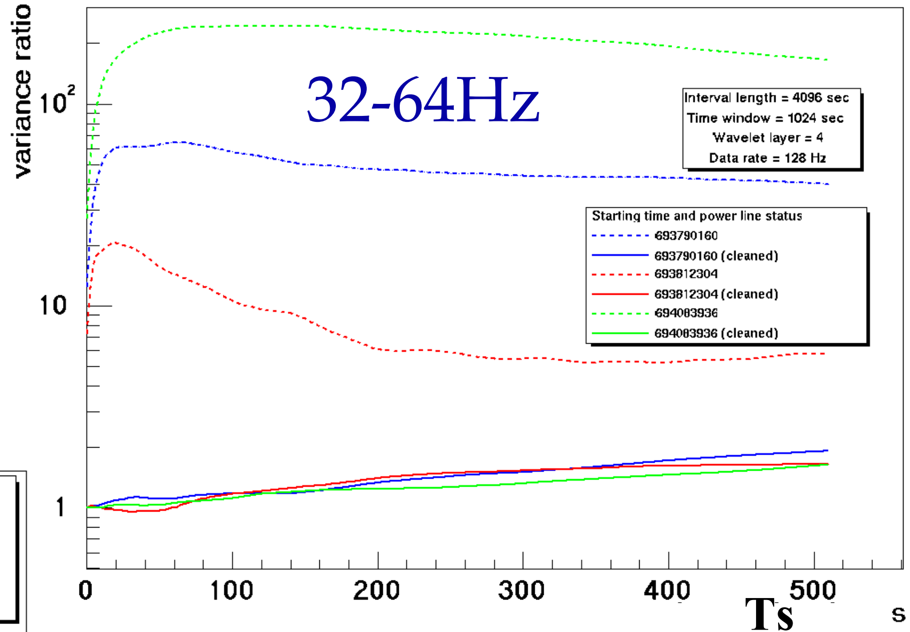


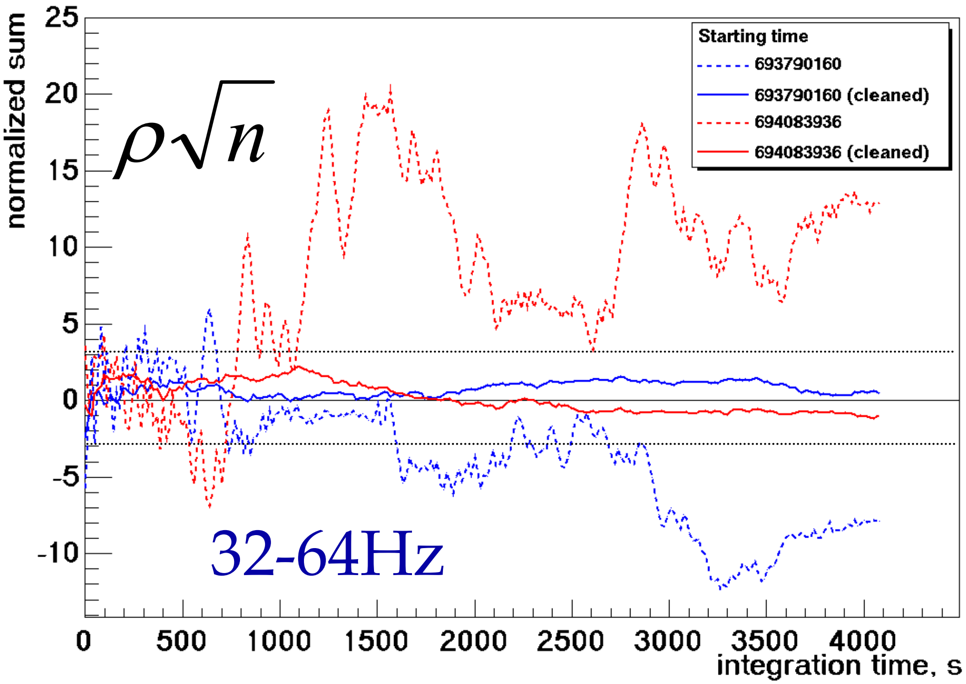




Data with Lines Removed

- QMLR method was used





$\pm 3/\sqrt{n}$

dots - before

solid - after

Wavelet Transforms

- time-frequency representation of data in wavelet domain
  - P_{mn} : n – scale (frequency) index, m – time index
- due to of locality of wavelet basis, wavelet layers are decimated time series.

$$x_L(t) = \sum_{k,l} p_{kl} \psi_{kl}(t)$$

$$x_H(t) = \sum_{n,m} q_{mn} \psi_{mn}(t)$$

- X-correlation

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

$$I_{kl,mn} = \int_0^T dt \int_0^T dt' \psi_{kl}(t') \psi_{mn}(t) Q(t - t', \Omega_L, \Omega_H)$$

  - Ψ_{nm} – basis of wavelet functions

Cross-Correlation in Wavelet Domain

- **x-correlation is a sum over wavelet layers**
  - ✓ **τ – time lag**
  - ✓ **k – wavelet layer number**
  - ✓ **N_k – number of samples in layer k**
  - ✓ **$r_k(\tau)$ – correlation coefficients as a function of lag time τ**

$$S = \sum_{k,\tau} N_k w_k(\tau) r_k(\tau)$$

- **$w_k(\tau)$ – optimal filter**

$$w_k(\tau) = \int_{-\infty}^{\infty} df \left|\psi_k(f)\right|^2 \left|f\right|^{-3} \frac{\Omega_{GW}(f)\cdot\gamma\left(f,\Omega_L,\Omega_H\right)}{P_L(f)/\sigma_k^L \cdot P_H(f)/\sigma_k^H} \exp\left(-j2\pi f\tau\right)$$

  - ✓ **Ψ_k – Fourier image of mother wavelet for layer k**
  - ✓ **γ – overlap reduction function**
  - ✓ **σ_k^L, σ_k^H - noise *rms* in wavelet domain for detector L (H)**
  - ✓ **$P(f)$ – noise PSD**

Sign Cross-Correlation

- **Sign x-correlation** $$S_s = \sum_{k,\tau} N_k \omega_k(\tau)\rho_k(\tau)$$ →
  - $\rho_k(\tau)$ **– sign correlation coefficients**
  - $\omega_k(\tau)$ **– optimal filter**
- **Variance of ρ** $$\mathrm{var}(\rho_k(\tau)) = \frac{1}{N_k} v_k(\tau)$$
- **optimal filter** $$\omega_k(\tau) = \varepsilon_k \frac{w_k(\tau)}{v_k(\tau)}$$ ←
  - ε_k **– sign correlation efficiency**
  - $v_k(\tau)$ **– contribution from correlated noise**
  - $w_k(\tau)$ **- optimal filter for linear correlation**

$$V_s = \sum_i N_i \omega_i^2 v_i$$

$$\bar{\rho}_i = \lambda_i \Omega$$

$$\mu = \Omega \sum_i N_i \omega_i \lambda_i,$$

$$\frac{\partial(SNR^2)}{\partial \omega_i} = \frac{\partial(\mu^2 / V_s)}{\partial \omega_i} = 0,$$

$$\omega_i = \lambda_i / v_i$$

$$\bar{r}_i = w_i \Omega$$

$$\lambda_i = w_i \frac{\bar{\rho}_i}{\bar{r}_i} = w_i \varepsilon_i$$

$$i = (k, \tau)$$

$\rho(\tau)$-signature of the correlation

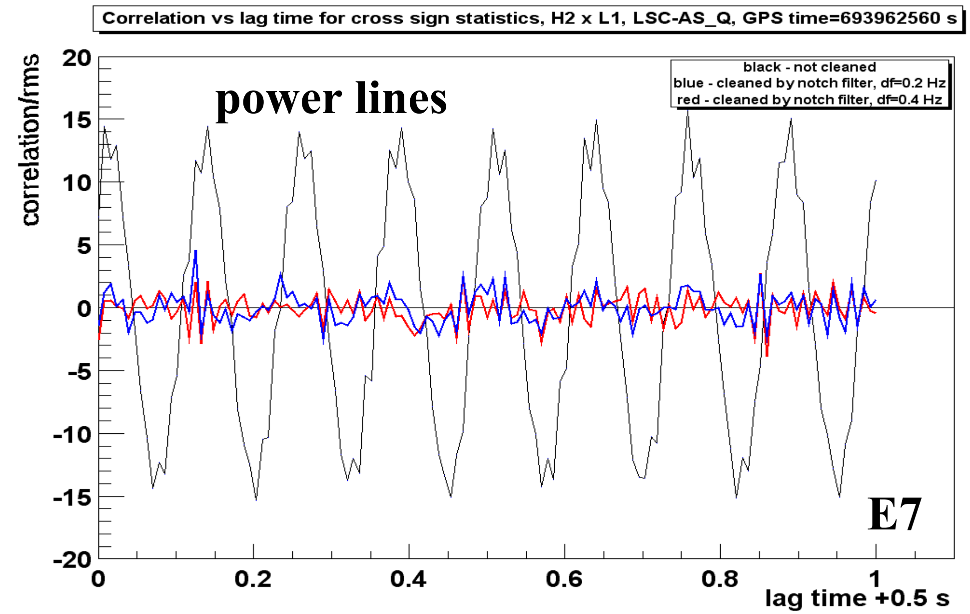
- **Likelihood function**

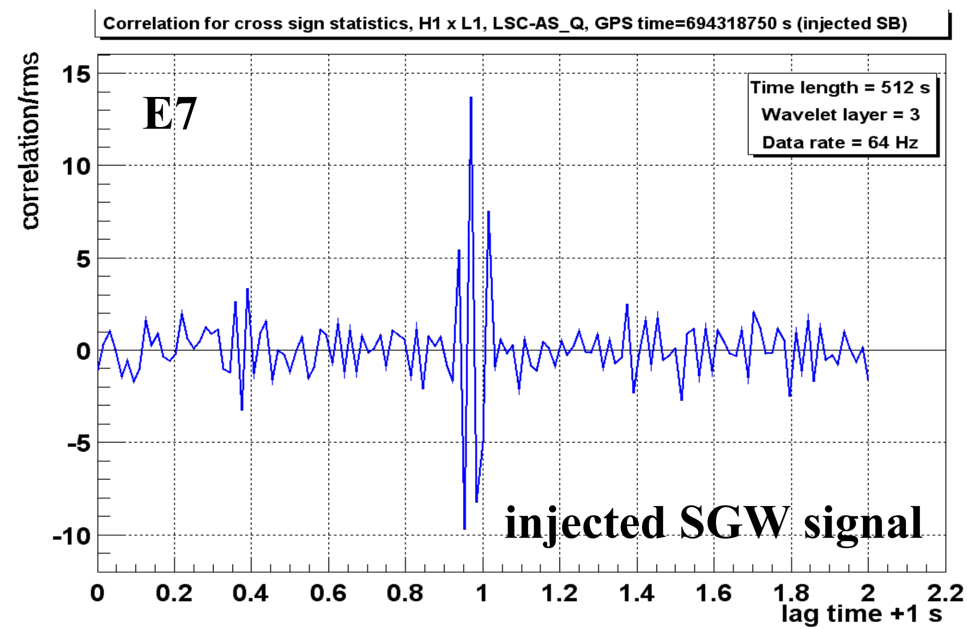
$$L = \sum_i \frac{N_i}{2\nu_i}\left(\rho_i - \overline{\rho}_i\right)^2 \qquad i = (k, \tau)$$

- **$\overline{\rho}_i$ depend on SGW and noise (parametric) models**

$$\overline{\rho}_i(\Omega_{GW}, \Omega_n, \alpha, \beta, \ldots.)$$

- **by measuring $\rho(\tau)$, SGW can be separated from correlated noise**





Robust Optimal Filter

- **Optimal filter**

$$\omega_k(\tau) = \frac{1}{\nu_k} \int_{-\infty}^{\infty} df \left|\psi_n(f)\right|^2 |f|^{-3} \frac{\Omega_{GW}(f)\cdot\gamma(f,\Omega_L,\Omega_H)}{A_L(f,k)\cdot A_H(f,k)} \exp(-j2\pi f\tau)$$

- **"noise amplitude"**

$$A_I(f,k) = \frac{P_I(f)}{\sigma_{kI}\sqrt{\varepsilon_k}}$$

- ***A(f)* is more robust then *P(f)* if noise is non-stationary**

- **test with simulated noise**
  - Gaussian noise (σ_g) + tail : total rms σ_n

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

Upper Limit

- **Cross-correlation & variance:**

$$S_s = \sum_{k,\tau} N_k \omega_k(\tau) \rho_k(\tau) \qquad V_s = \sum_{k,\tau} N_k \omega_k^2(\tau) v_k(\tau) \qquad \omega_k(\tau) = \varepsilon_k \frac{w_k(\tau)}{v_k(\tau)}$$

- **x-correlation expectation value:** $$\mu = \Omega \sum_{k,\tau} N_k \omega_k^2(\tau) v_k(\tau) = \Omega V_s$$

- **signal to noise ratio:** $$SNR = \Omega \sqrt{\sum_{k,\tau} N_k \omega_k^2(\tau) v_k(\tau)} = \Omega \sqrt{V_s}$$

- **confidence level:** $$CL = \frac{1}{2} erf\left(\frac{S_s}{\sqrt{V_s}}\right) \rightarrow \tilde{S}_s(95CL)$$

- **upper limit:** $$\tilde{\Omega} = \frac{\tilde{S}_s(95CL)}{V_s}$$

Summary

- A robust correlation test with treatment of correlated noise is described. It allows:
  - **calculate x-correlation distribution if noise is not Gaussian**
  - **use a simple model of correlated noise**
  - **work for non-stationary noise**
- suggested method offers a good tool to estimate contribution from correlated noise.
  - **On E7 data it is shown how the noise affects the x-correlation.**
- we suggest to use sign x-correlation as a complementary method for setting SGW upper limit
  - **a very useful cross-check**
  - **uses $\rho_k(\tau)$ – signature of the cross-correlation**