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

S.Klimenko, University of Florida

Outline

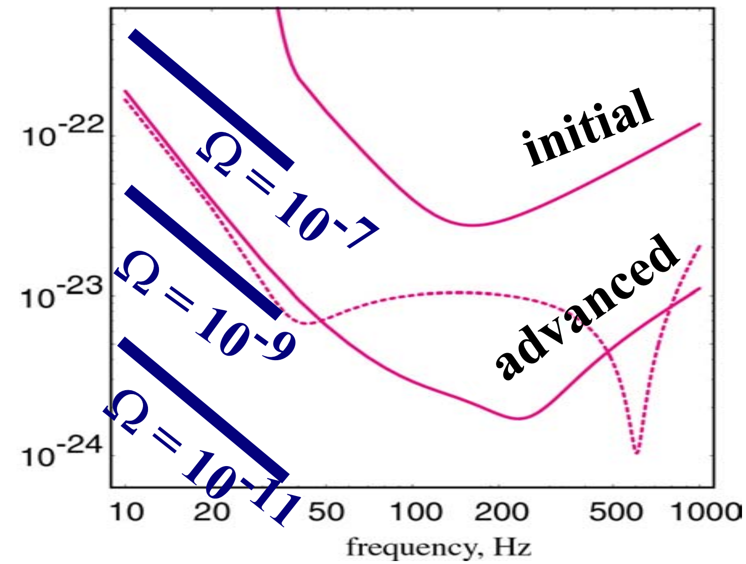
- Introduction
- Optimal cross-correlation
- Correlation tests
- Cross-correlation in wavelet domain
- Correlated noise
- Conclusion

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 detectors output** $s_L(t)$ ***and*** $s_H(t)$

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

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

Optimal Cross-Correlation

- **x-correlation in Fourier domain** **(B.Allen,J.Romano gr- qc/ 9604033 v3 30 Sep 96)**

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

  - **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 *distribution of S* if noise is not Gaussian?**
  - **What to do if noise is not stationary?**
  - **What to do if *S* is affected by correlated (©) noise?**

Correlation Tests

- **linear correlation test**

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

- correlation coefficient

  - 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 correlation test**
  - non-parametric: exactly known ***r*** distribution
  - effective but CPU inefficient for large data sets

Sign Correlation Test

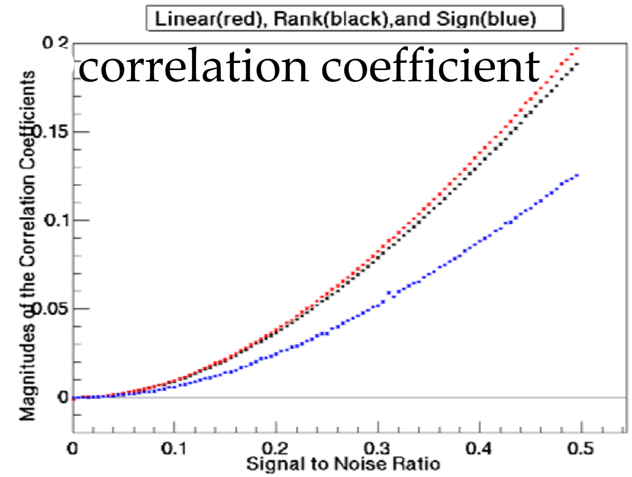
- Sign transform: $$u_i = sign(x_i - \hat{x})$$
  - $\hat{x}$ - **median of x**
- Sign statistics: $$s_i = sign(x_i - \hat{x}) \cdot sign(y_i - \hat{y})$$
- Correlation coefficient γ: $$\gamma = mean(s_i)$$
- Distribution of γ (n – number of samples):
  - **Gaussian (large n):** $$P(n,\gamma) \approx \sqrt{\frac{n}{2\pi}} \cdot \exp\left(-\frac{n\gamma^2}{2}\right)$$
- very robust:
  - error from $\hat{x}$ and $\hat{y}$ $\sim 2/n^2$, much less then $\boldsymbol{var(\gamma)=1/n}$ for large n

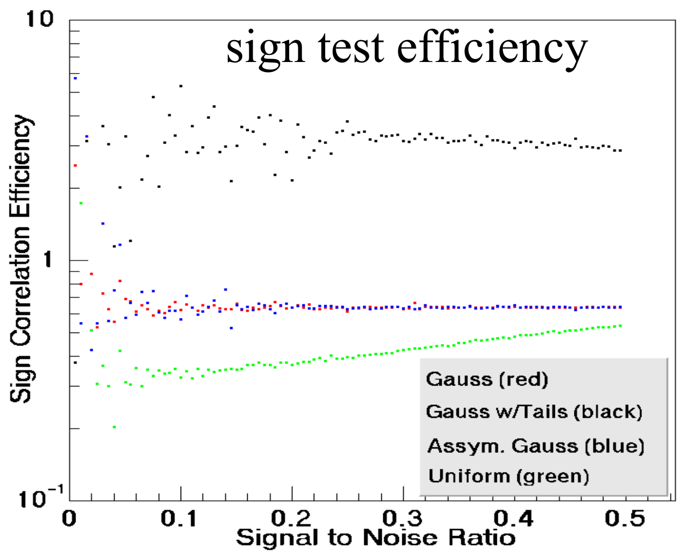
Comparison of Correlation Tests

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

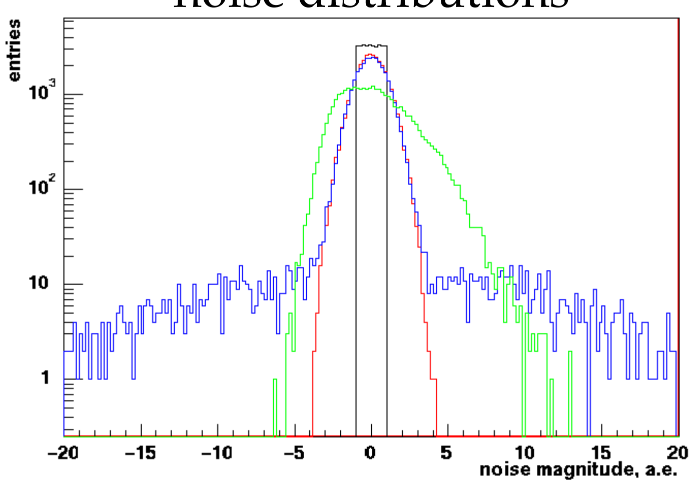
$$x = n_x + g, \qquad y = n_y + g$$

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





noise distributions



S.Klimenko

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 (similar to windowed FT).
- X-correlation

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

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

  - Ψ_{nm} – basis of wavelet functions

Cross-Correlation in Wavelet Domain

- x-correlation is a sum over wavelet layers

$$S=\sum_{n,\tau} w_n(\tau) r_n(\tau)$$

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

  - τ – time lag
  - n – wavelet layer number
  - N_n – number of samples in layer n
  - $r_n(\tau)$ – correlation coefficients as a function of leg time τ
  - $w_n(\tau)$ – optimal weight
    - ✓ Ψ_n – Fourier image of mother wavelet for layer n
    - ✓ σ_{nL},σ_{nH} - noise *rms* in wavelet domain for detector L (H)
- equivalent to *S* calculated in Time & Fourier domains

Sign X-Correlation in Wavelet Domain

- replace $r_n(\tau)$ with $\gamma_n(\tau)$ - sign correlation coefficients
- To keep the weights optimal, take into account the sign correlation efficiency ε_n

$$S_s = \sum_{n,\tau} \tilde{w}_n(\tau)\gamma_n(\tau), \qquad \tilde{w}_n(\tau) = w_n(\tau)\varepsilon_n$$

- $\gamma_n(\tau)$ are normally distributed with variance $1/N_n$
  - then the x-correlation variance is: $\mathrm{var}(S_s) = \sum_{n,\tau} \frac{1}{N_n} \tilde{w}_n^2(\tau)$
- What we gain/loose
  - sign test is less efficient (65%) when data is Gaussian:
  - if data is not Gaussian
    - ✓ the sign test can be more efficient
    - ✓ gain confidence in calculation of S distribution

Robust Spectral Amplitude

- **Optimal weight**

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

- **“noise amplitude”**

$$A_I(f) = \frac{P_I(f)}{\sigma_{nI}\sqrt{\varepsilon_n}}$$

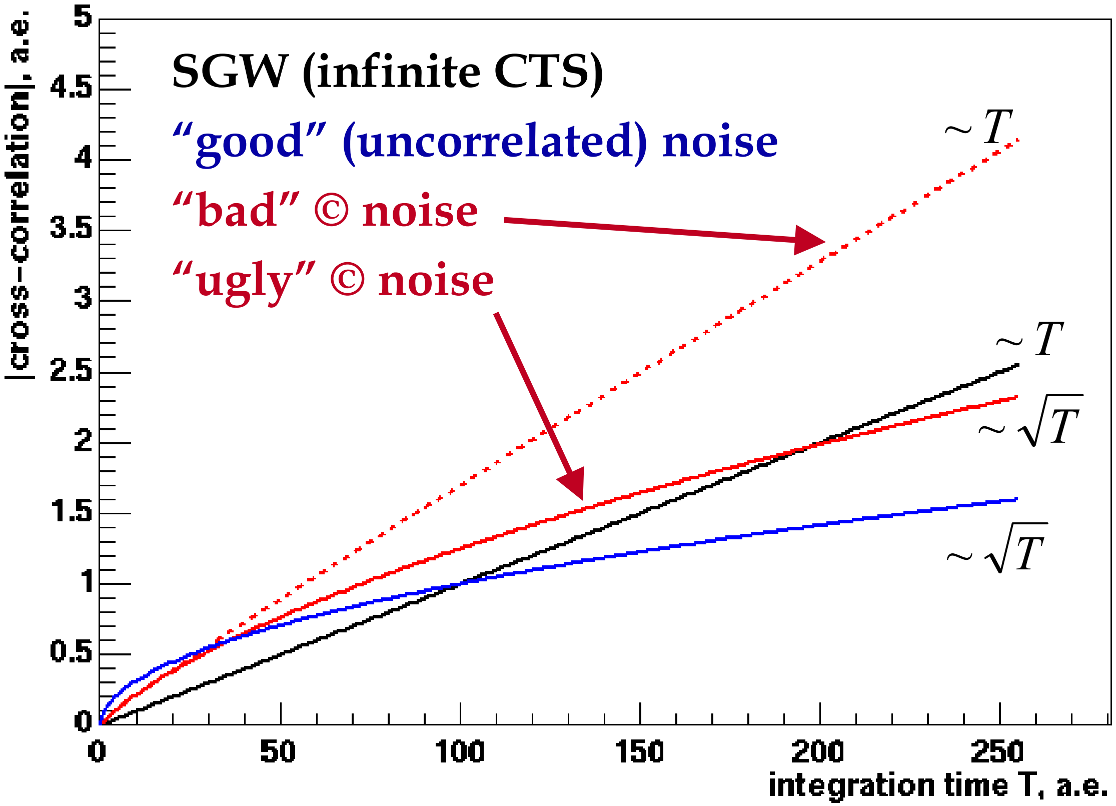
- ***A(f)* is more robust then *P(f)* if noise is non-stationary**

- **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

© Noise

$$C = \langle h_L + n_L, h_H + n_H \rangle = \langle h_L, h_H \rangle + \langle n_L, n_H \rangle$$

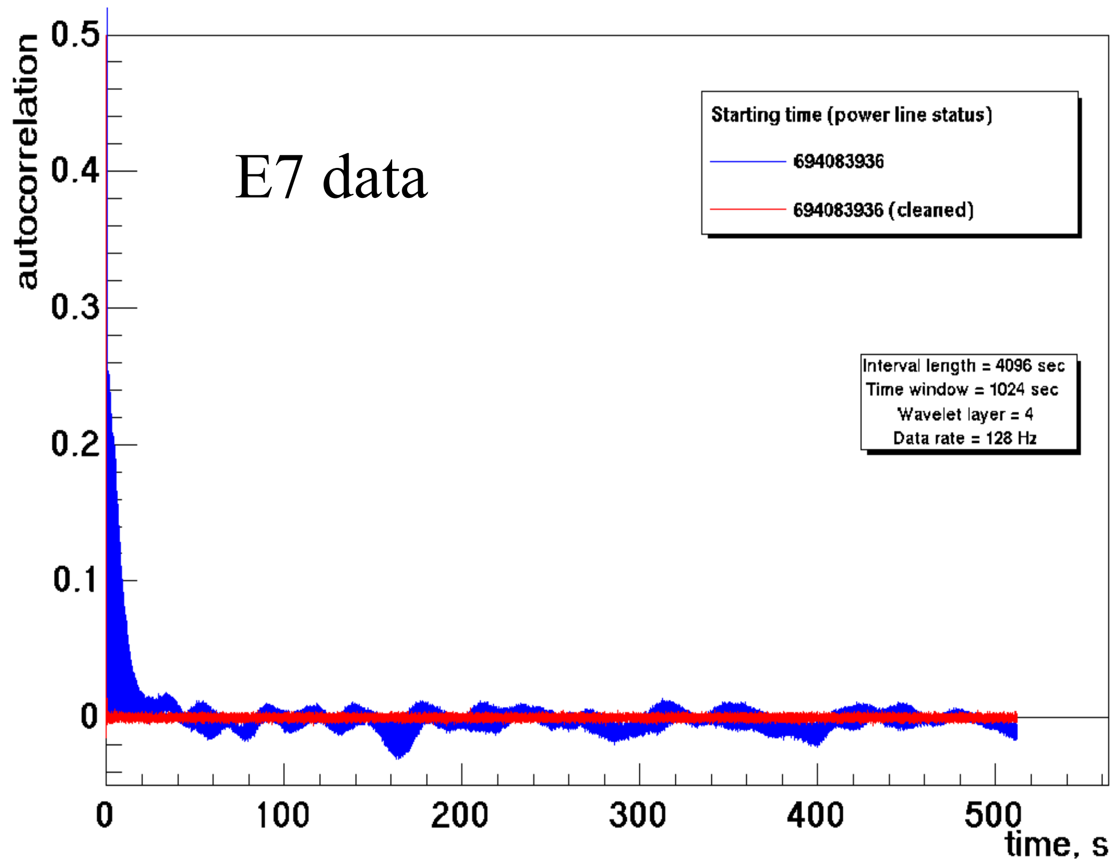


- **remove "bad" © noise (likely to be data processing artifacts)**
- **How to deal with "ugly" © noise?**

Autocorrelation Function

- sign statistics $s(t)=\{u_x u_y\}$
- $a(t)$ - autocorrelation function of $s(t)$
  - **a measure of correlated noise.**

X-correlation of
L1:AS_Q & H2:AS_Q
in wavelet domain:
32-64 Hz band



Noise Model in SCT

- **uncorrelated noise**
  - **autocorrelation function:** $a(0) = 1,\ a(\tau \geq \Delta t) = 0$
  - **null hypothesis:** ***data sets are not correlated***
  - **variance:** $\mathrm{var}_0(\gamma) = \frac{1}{n}$
- **correlated noise 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:** $\mathrm{var}_{Ts}(\gamma) = \frac{1}{n} R,$ $\quad R = 1 + \sum_{m=1}^{T_s/\Delta t} (n-m) a_n(m\Delta t)$
- **SCT allows calculate *var(γ)*, depending on the noise model.**

R

- **variance ratio**

$$R = \frac{\mathrm{var}_{Ts}(\gamma)}{\mathrm{var}_0(\gamma)}$$

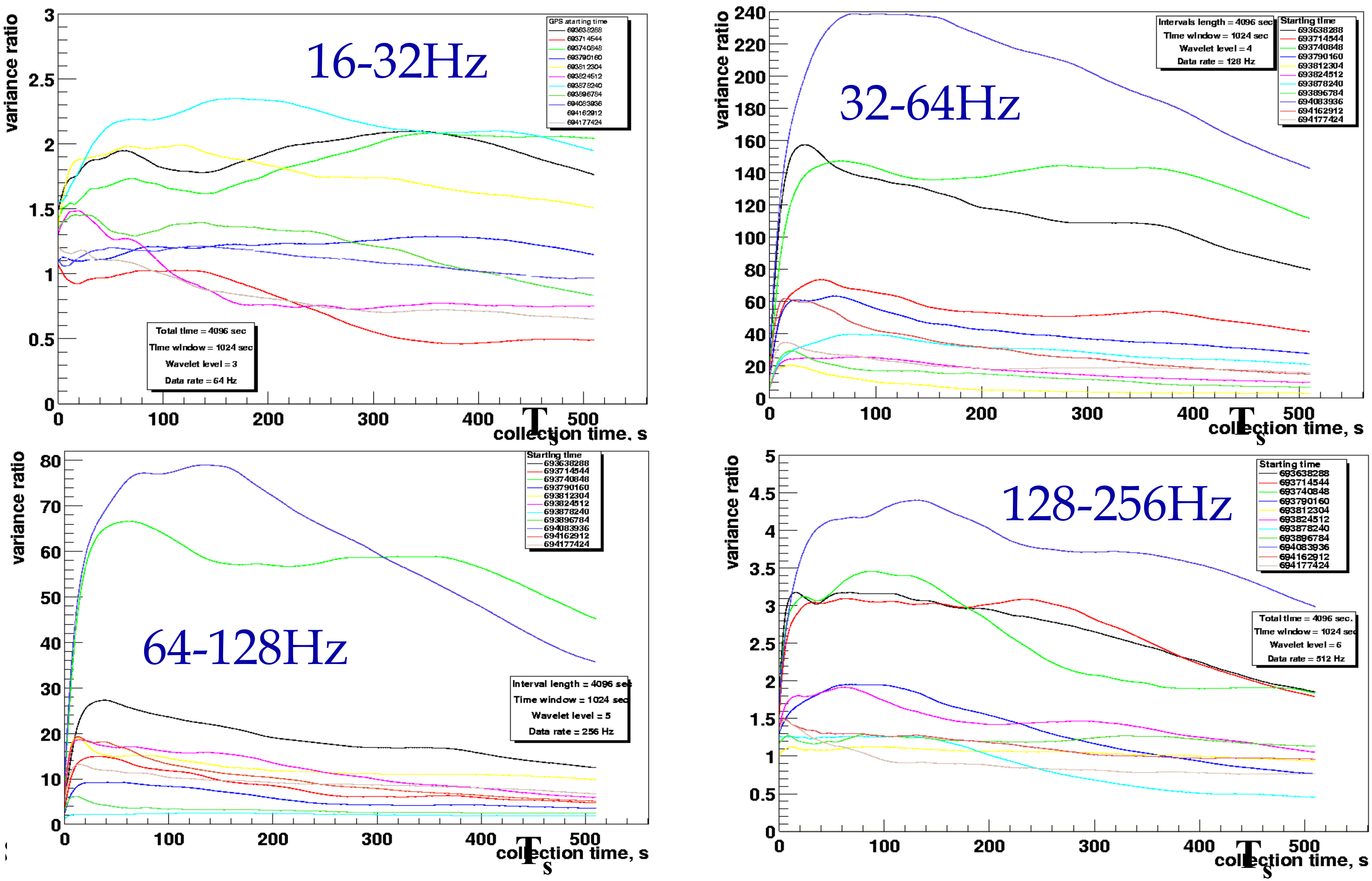
  - **R is a measure of © noise, or quality of data.**
  - **R times more data needed to reach same CL as for uncorrelated noise.**
  - **If R is too large, the noise should be removed, if possible**

- **residual correlated noise is handled by**
  - **reduced correlation coefficient:** $\gamma' = \gamma R^{-1}$
    - ✓ **normally distributed with variance** ***1/n***

- **x-correlation in wavelet domain:** $S_s = \sum_{n,\tau} \tilde{w}_n(\tau)\gamma'_n(\tau),$

Variance Ratio (T_s)

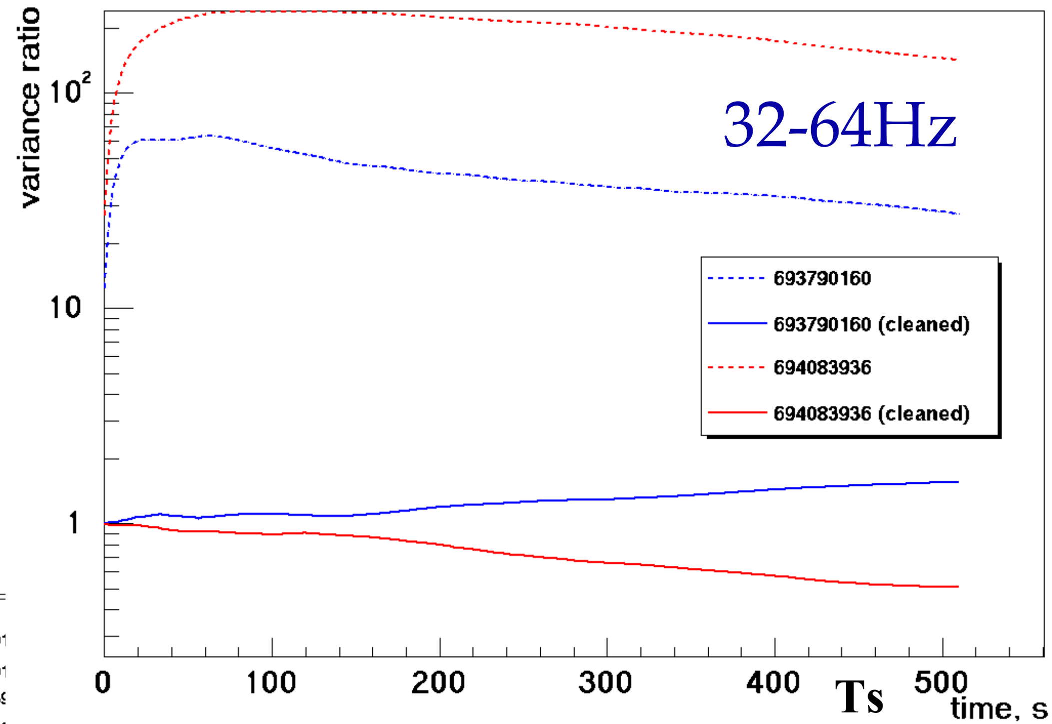
- **L1xH2: 11 data segments 4096 sec each (total 12.5 h of E7 data)**

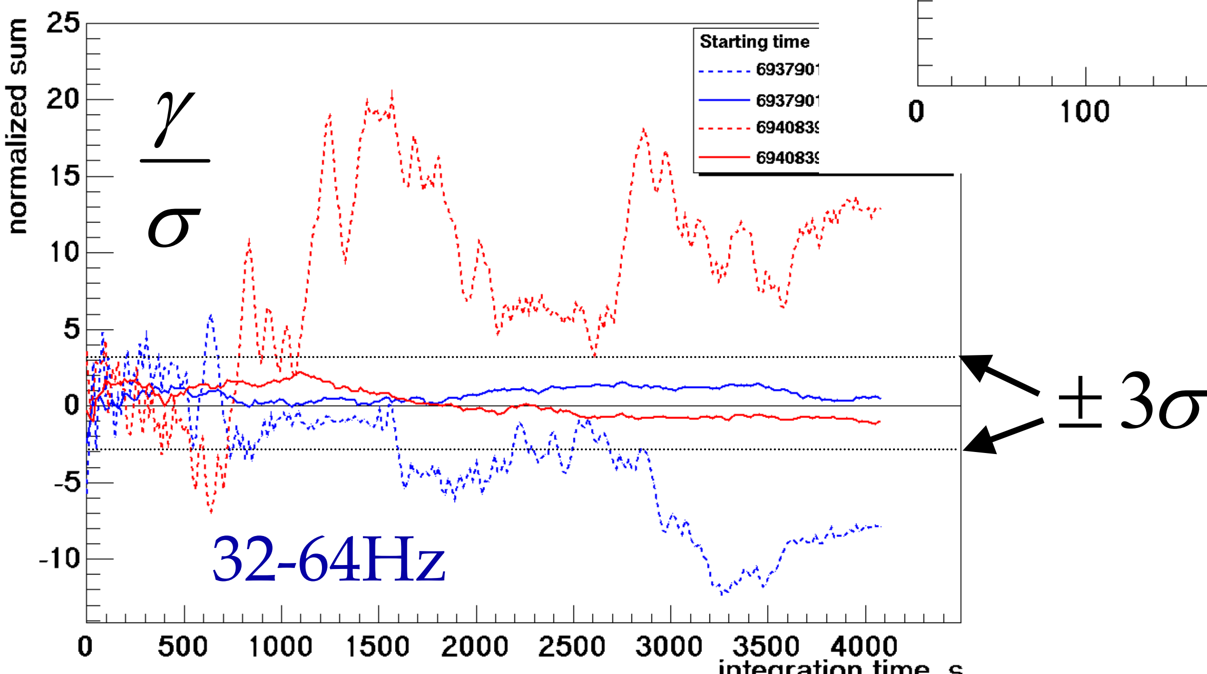


Data with Lines Removed

- QMLR method was used

σ – rms for uncorrelated noise





dots - before

solid - after

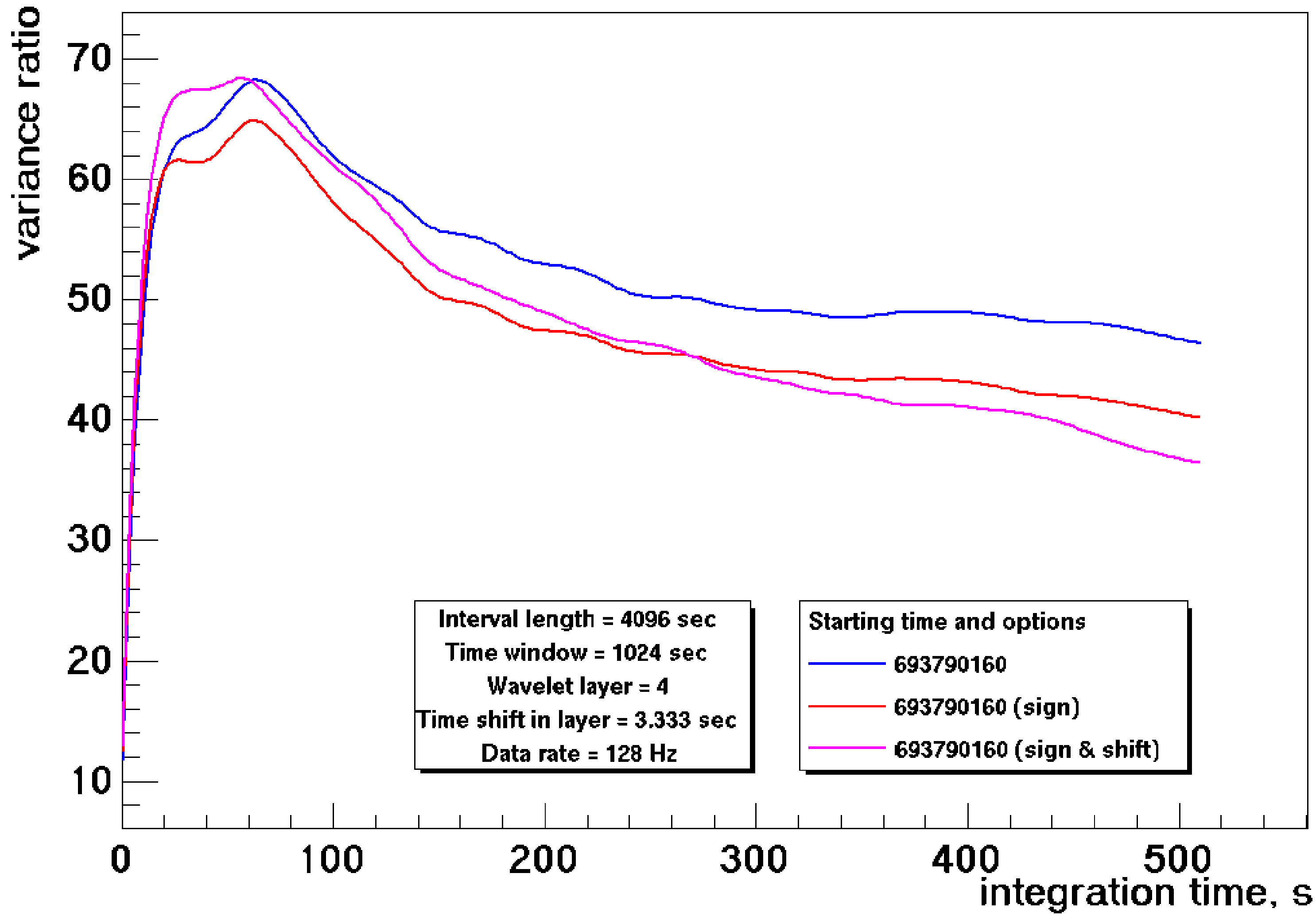
Setting Upper Limit

1. ***correlation coefficients*** γ – measured
2. ***variance of*** γ – calculated for given model of © noise
   - **© noise can be estimated from data if Ts<<T**
3. ***optimal coefficients w*** – calculated for given SGW model.
   - **sign correlation efficiency ε – estimated from simulation.**
4. as a result of 1,2,3 calculate x-correlation $$S_s = \sum_{n,\tau} \tilde{w}_n(\tau)\gamma'_n(\tau),$$
5. find from simulation the dependence $S(\Omega_{sim})$ ($\sim a\Omega_{sim}$)
6. Set upper limit by calculating confidence belts.

Summary

- A robust correlation test with treatment of © noise is described. It allows:
  - **calculate x-correlation distribution if noise is not Gaussian**
  - **work with non-stationary noise**
  - **use a simple model of correlated noise**
- suggested method offers a good tool to estimate contribution from © noise.
  - **On E7 data it is shown how © noise affects the x-correlation.**
- we suggest to use sign x-correlation as a complementary method for setting SGW upper limit
  - **very simple and CPU efficient**

R



variance ratio
70
60
50
40
30
20
10
0
100
200
300
400
500
integration time, s
Interval length = 4096 sec
Time window = 1024 sec
Wavelet layer = 4
Time shift in layer = 3.333 sec
Data rate = 128 Hz
Starting time and options
693790160
693790160 (sign)
693790160 (sign & shift)