Interferometric Data Modeling: Issues in realistic data generation.

Soma Mukherjee

CGWA

Dept. of Physics and Astronomy

University of Texas Brownsville

LIGO Tech Doc #LIGO-G040364-00-Z

Why do we need a model ?

- Astrophysical searches estimate efficiency from playground data.
- Externally triggered search (Gamma Ray Bursts with GW) – how representative is the ’off-source’ segment of the ‘on-source’ one ?

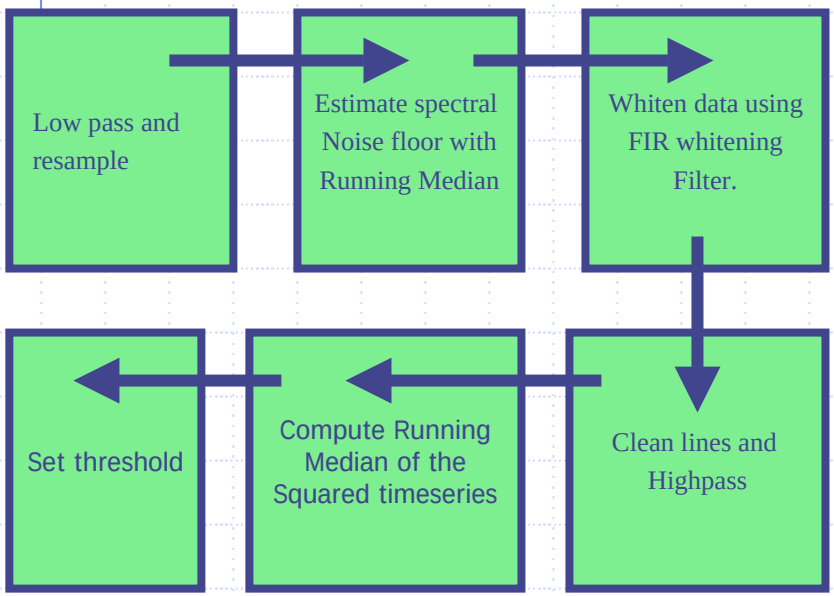
Data Modeling: Basis

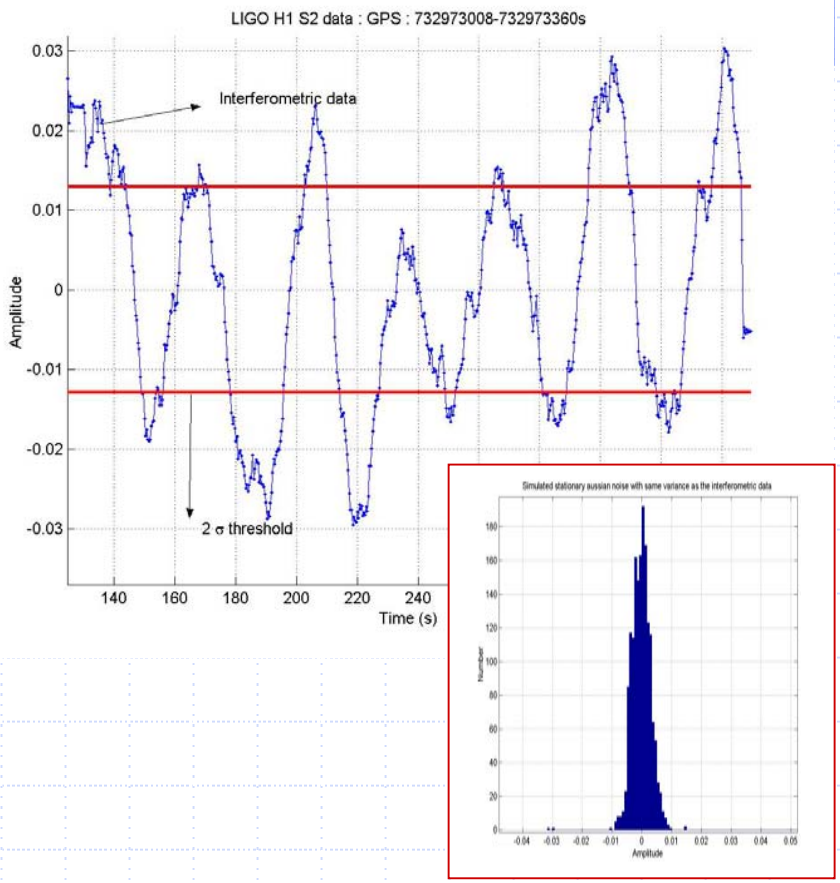
- Interferometric data has three components : lines, transients and noise floor.
- As a first approximation, the three components are independent and appear additively.
  - Physically different sources for each
- Basic idea is to split a channel into these components with mutual exclusion
  - Classify Transients, fit ARMA models to line amplitude and phase modulation, ARMA models for noise floor rms

Issue I :Slowly drifting noise floor

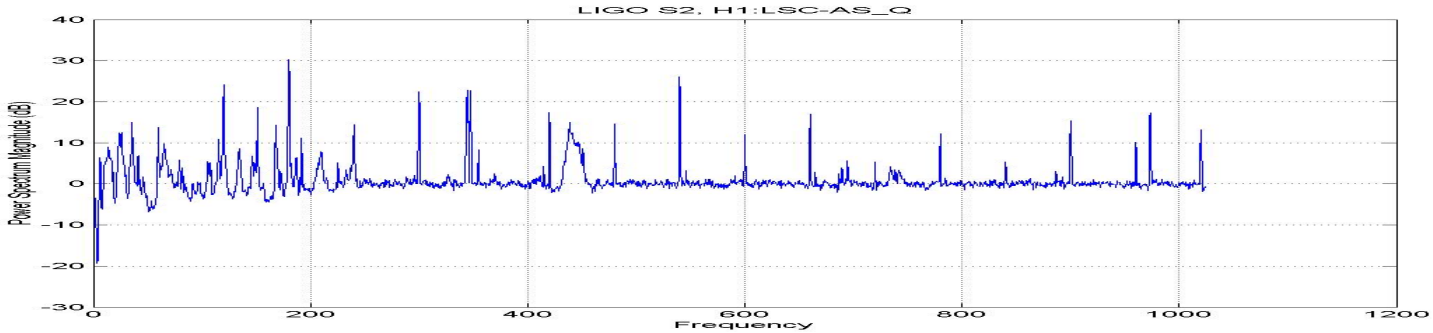
- Data from present generation of interferometers is non-stationary.

Results obtained by running MNFT[1] on a stretch of LIGO S2 data.[[1]Mukherjee, CQG, 2003]

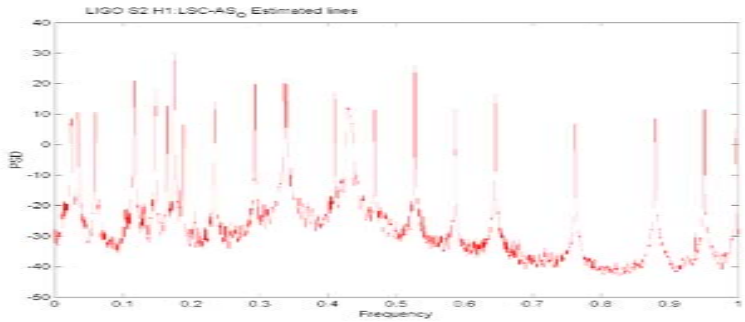


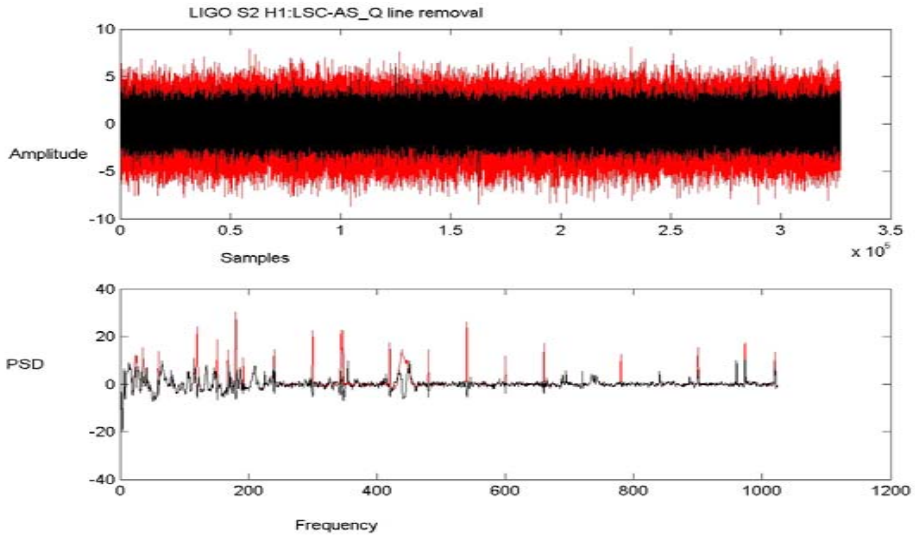


Issue II : Modeling ALL lines



MBLT[1] : Non-parametric Line estimation. [[1] Mohanty CQG, 2001]

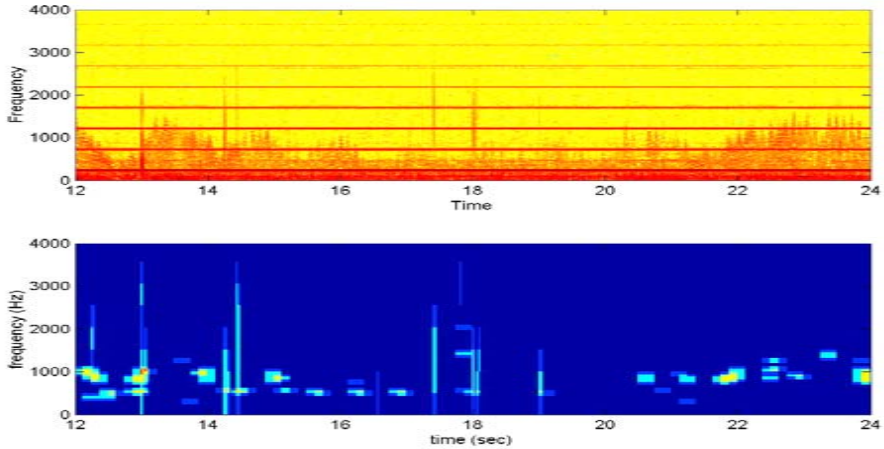


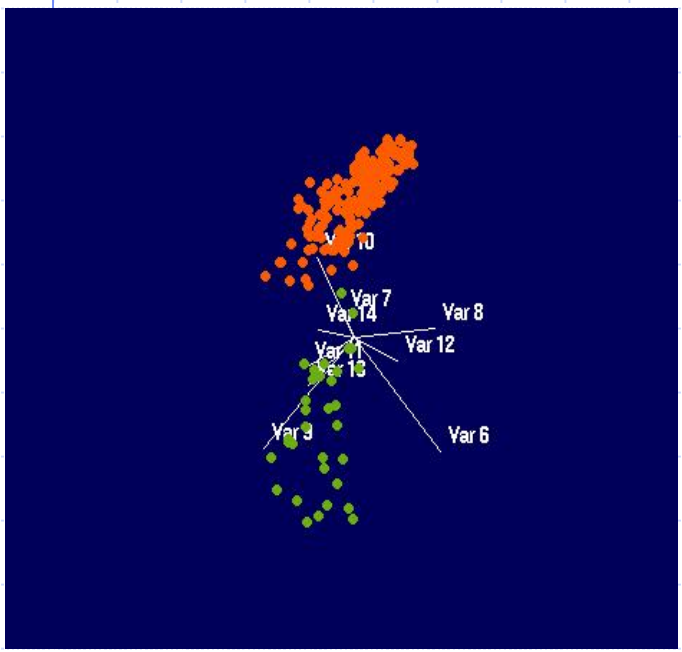


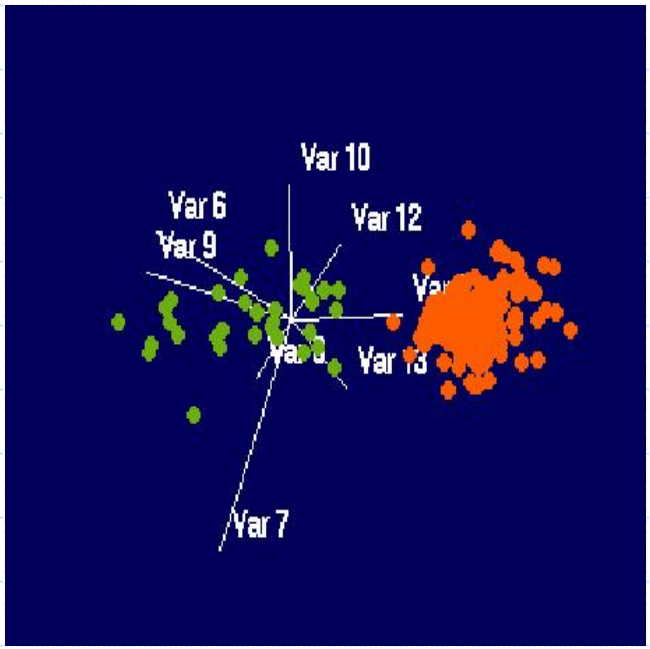
Issue III : Transient Classification

Non-parametric change point detector.

KSCD : Kolmogorv-Smirnov test based Change point Detector Mohanty, GWDAW (2002); PSDCD, Mohanty, PRD (2000);







12 top wavelet coefficients of data surrounding each KSCD trigger. *Visualized using GGobi.* (Preliminary).

Mukherjee, 2003, Amaldi, Pisa

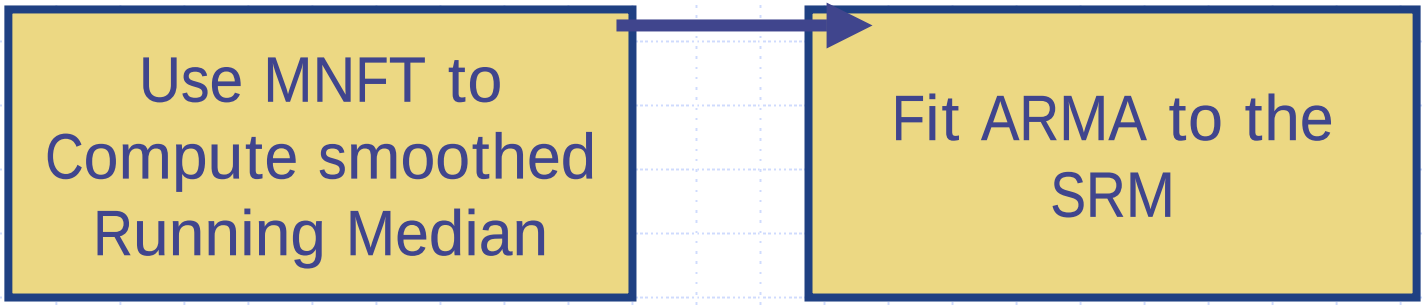
MNFT outline:

Algorithm:

1. Lowpass and resample given timeseries *x(k).*
2. Construct FIR filter that whitens the noise floor. Resulting timeseries : *w(k)*
3. Remove lines using notch filter. Cleaned timeseries : *c(k)*
4. Track variation in second moment of *c(k)* using Running Median and apply smoothing (SRM).
5. Obtain significance levels of the sampling distribution via Monte Carlo simulations.

Model Noise Generation

- Model Noise Floor (low order ARMA).



- Estimate lines using MBLT , ARMA model amplitude and phase, add reconstructed lines to synthetic data .
- Add transients.

ARMA (*p,q*)

A(T) y(t) = C(T) e(t)

Y(t) : Output

e(t) : White noise

C(T)/A(T) : Transfer function

T: Time shift operator

A and *C* : Polynomials

How faithful is the model ?

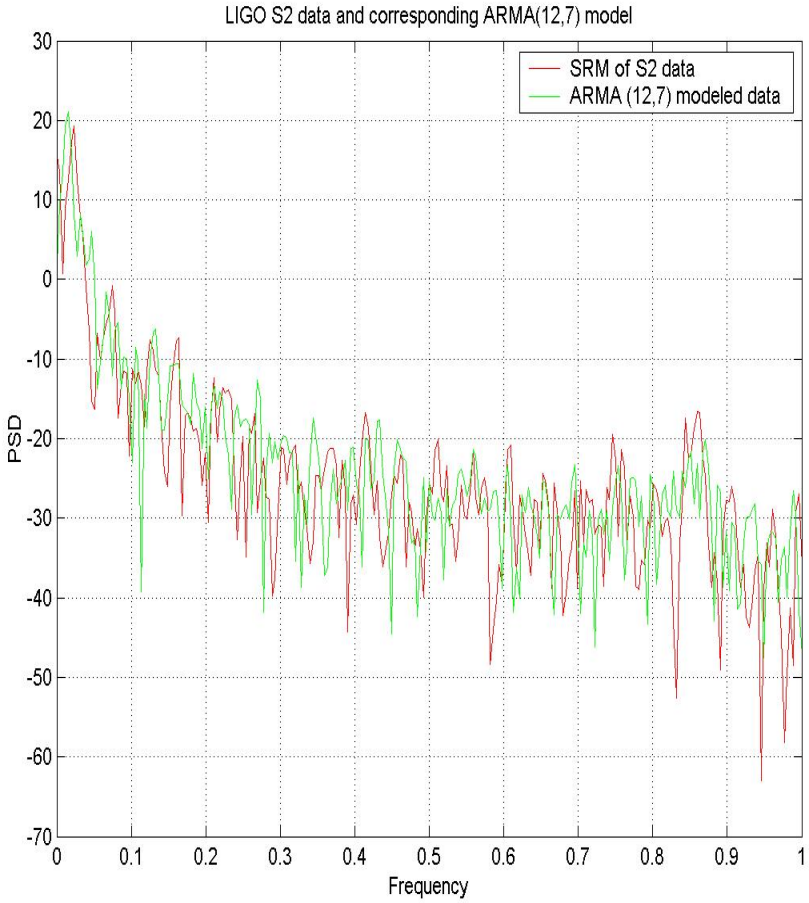
Apply statistical tests of hypothesis

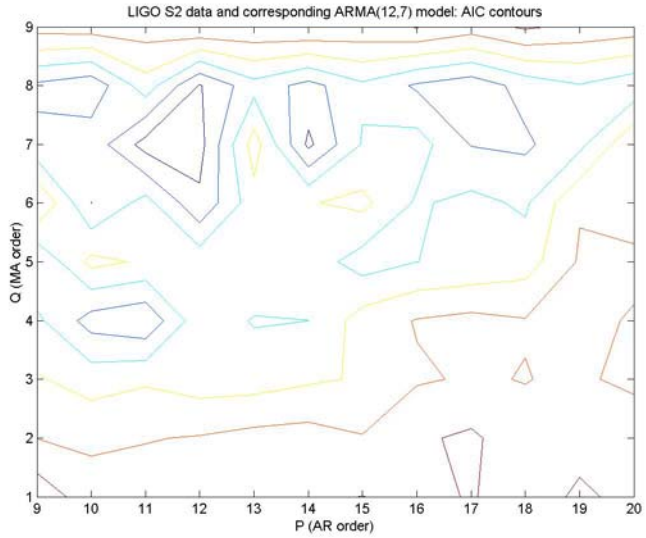
- Kolmogorov-Smirnov

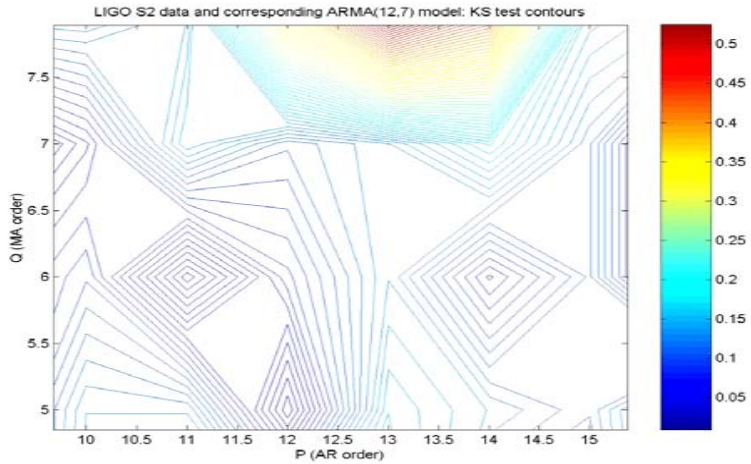
- Akaike Information criterion (AIC)

$$I_{akaike}(p,q) = \ln \sigma^{2}_{p,q} + 2(p+q)/N$$

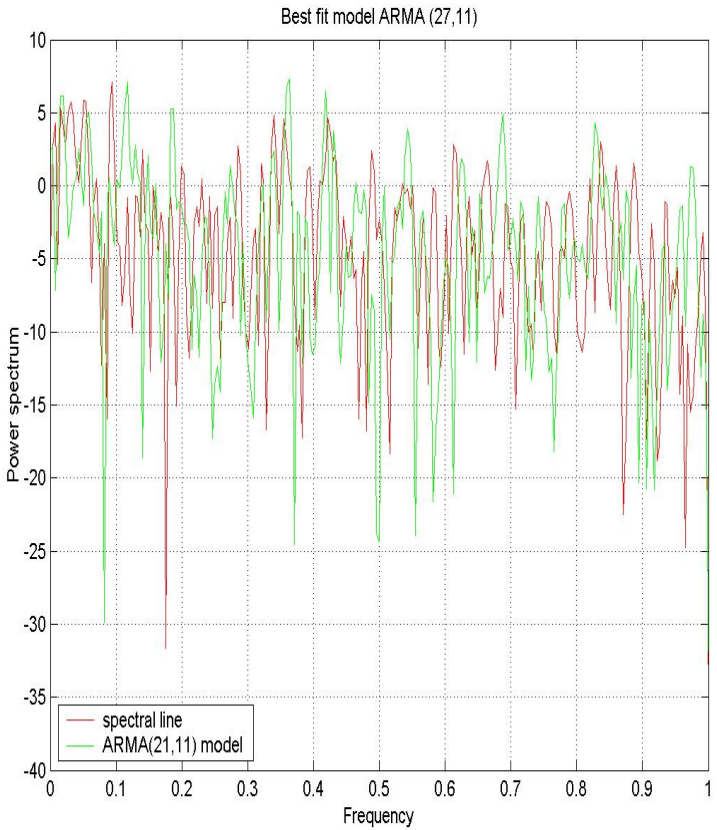
Result I : Noise floor model – ARMA (12,7)

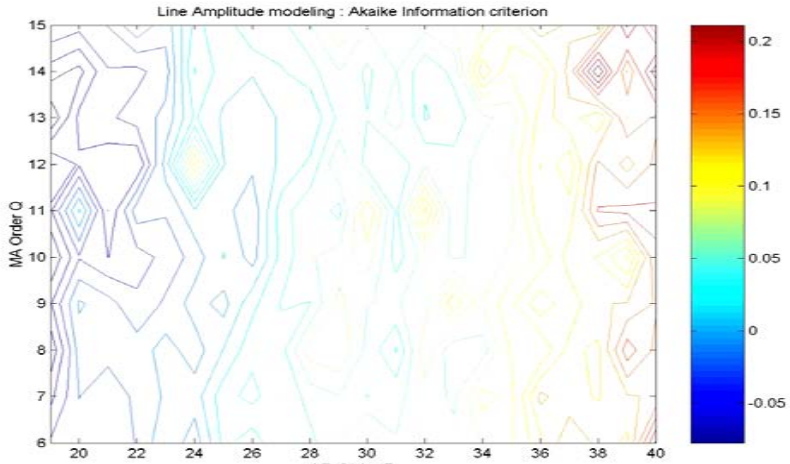


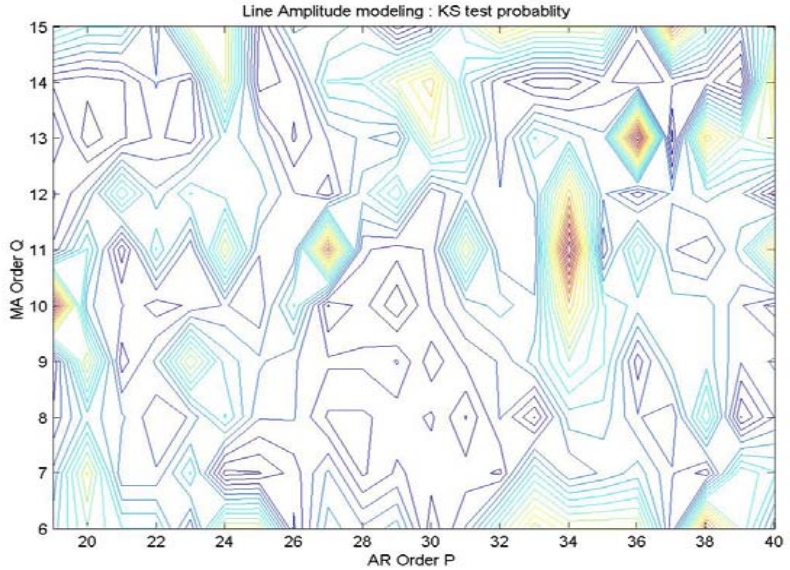




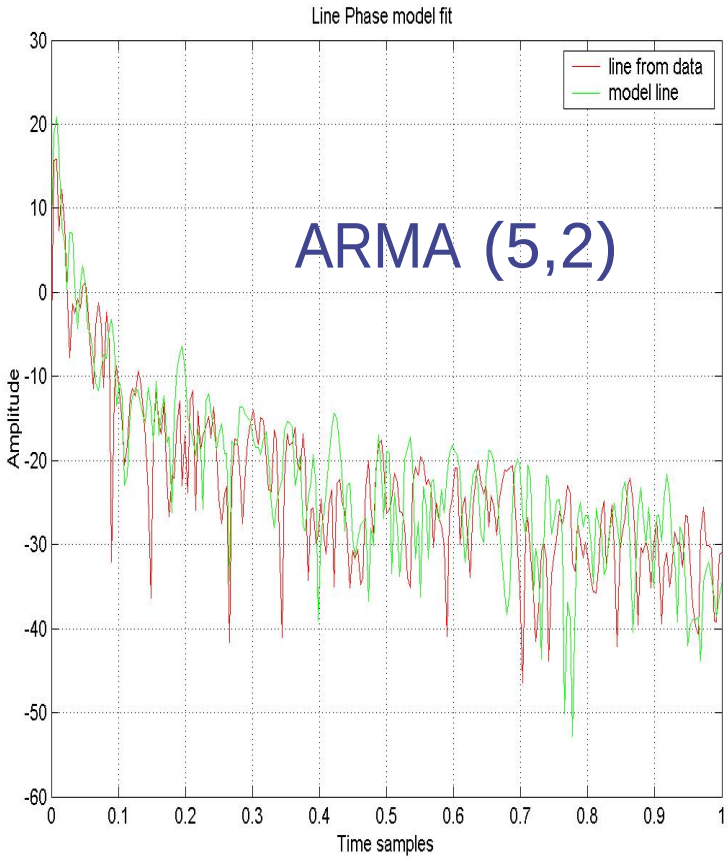
Result II : Line Amplitude - ARMA (27,11)

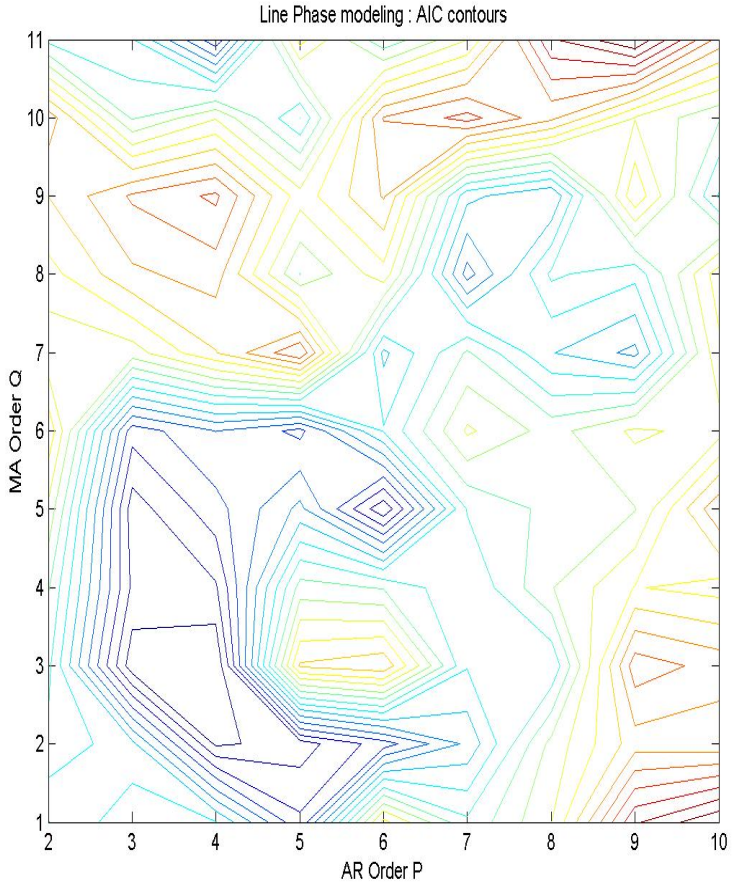






Line model : Phase – ARMA (5,2)





Plans

- Applicable to band limited data.
- Use as an ‘infinite’ playground for astrophysical searches.
- Gives a handle on non-stationarity and hence testing the robustness of the search algorithm.
- Allows us to do ‘controlled tests’.
- Signal injection and efficiency estimation