Data Analysis Challenge for the Ground-Based Interferometric GW Detectors

Linqing Wen

University of Western Australia, Perth

(Graphic source: NASA)

Outline

- Background
    - LIGO/PTA experiment: similarity and difference
- Interferometric detector (LIGO, VIRGO) data analysis
    - Individual detection methods
        - Inspiral source
    - Coherent network analysis/source localization
- Discussion
    - LIGO method to PTA data?

From the interferometer camp…

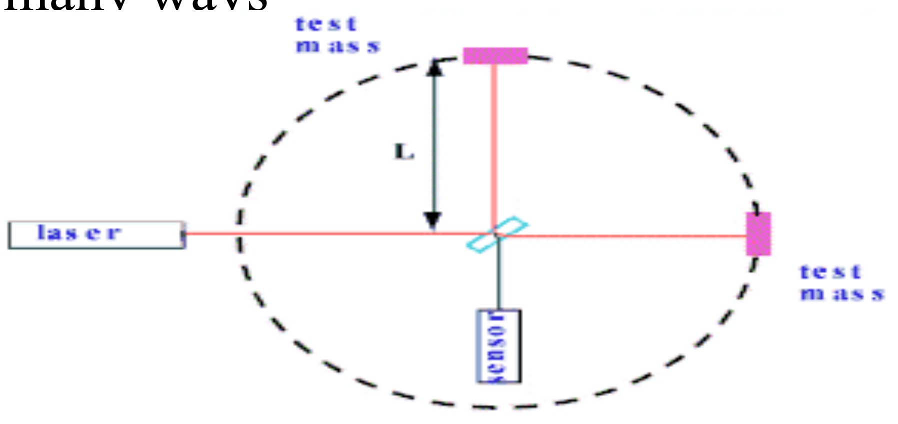
- Stas Babak: F- statistics, continuous source, LISA
- Sam Finn: matched filtering, individual source
- Xavier Siemens: stochastic background
- Alberto Sesana: SMBH mergers, Monte Carlo,…

Gravitational Waves

- GW caused by acceleration of mass quadrupole moment
  - e.g., binary coalescence
- Transverse wave
- Two wave polarizations (h_+, h_x)
- First direct detection yet to be done
- Ultimate test of Einstein's General Relativity

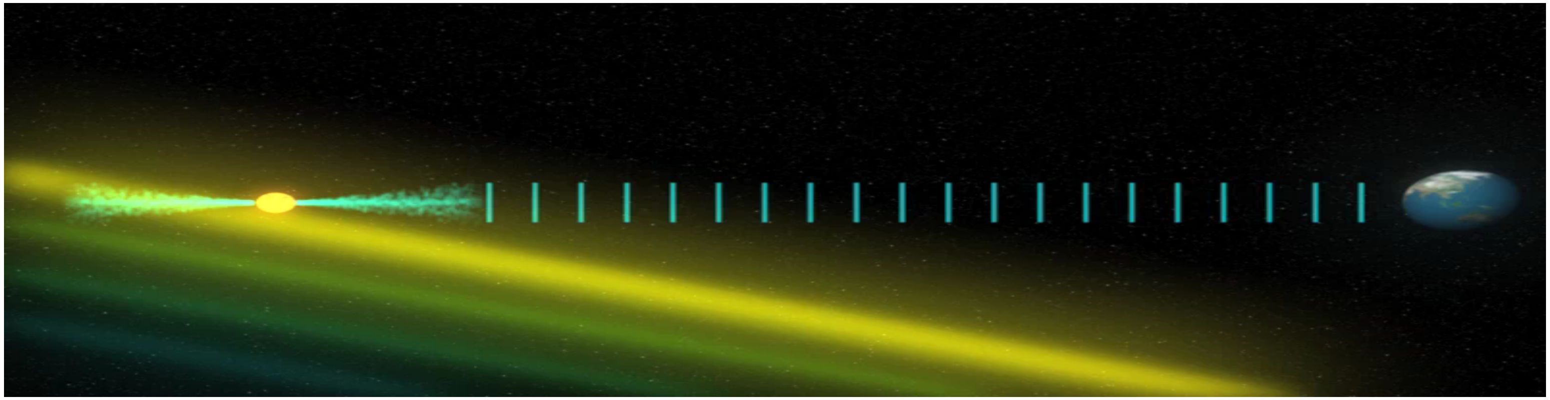
The Way We See Things…

- GW pushes freely floating object apart and together
- Laser interferometer is used to measure relative length change
  - Laser light travels in vacuum
- L-shaped Michelson interferometer with long arms
  - LIGO L = 4 km
  - f~10-2000 Hz
  - $\lambda_{GW} >> L$
  - Output: time-varying light intensity ~ h ~ $\Delta L/L$
    - $I(t) \sim (\Delta\phi)^2 \sim (4\omega_0^2 L \Delta L)\, h(t) + O(h^2)$
- Noise can be characterised in many ways
  - 100s' of environmental cl



Detecting GWs using Earth-Pulsar Detector

- GW changes travel time of radio pulses
- Measuring timing residual induced by a GW
- Need to detect detector behaviour at the same time
  - radio source behaviour ("laser") – binary motion, pulse evolution
  - Pulses go through interstellar medium
  - Poor knowledge on the pulsar distance
- Sensitivity
  - $f \sim 10^{-8}$ Hz
  - $\lambda_{GW} \ll L$
  - Typical $\Delta\tau \sim \lambda_{GW}\, h / c$ (~ 100 ns @ h~1e-15, $\lambda \sim$ pc)



Movie credit: http://www.gwastro.org/for%20scientists/visualization-pulsar-timing-and-gravitational-wave-detection

Interferometer GW Detectors

Laser Interferometer Gravitational-Wave Observatory (LIGO)

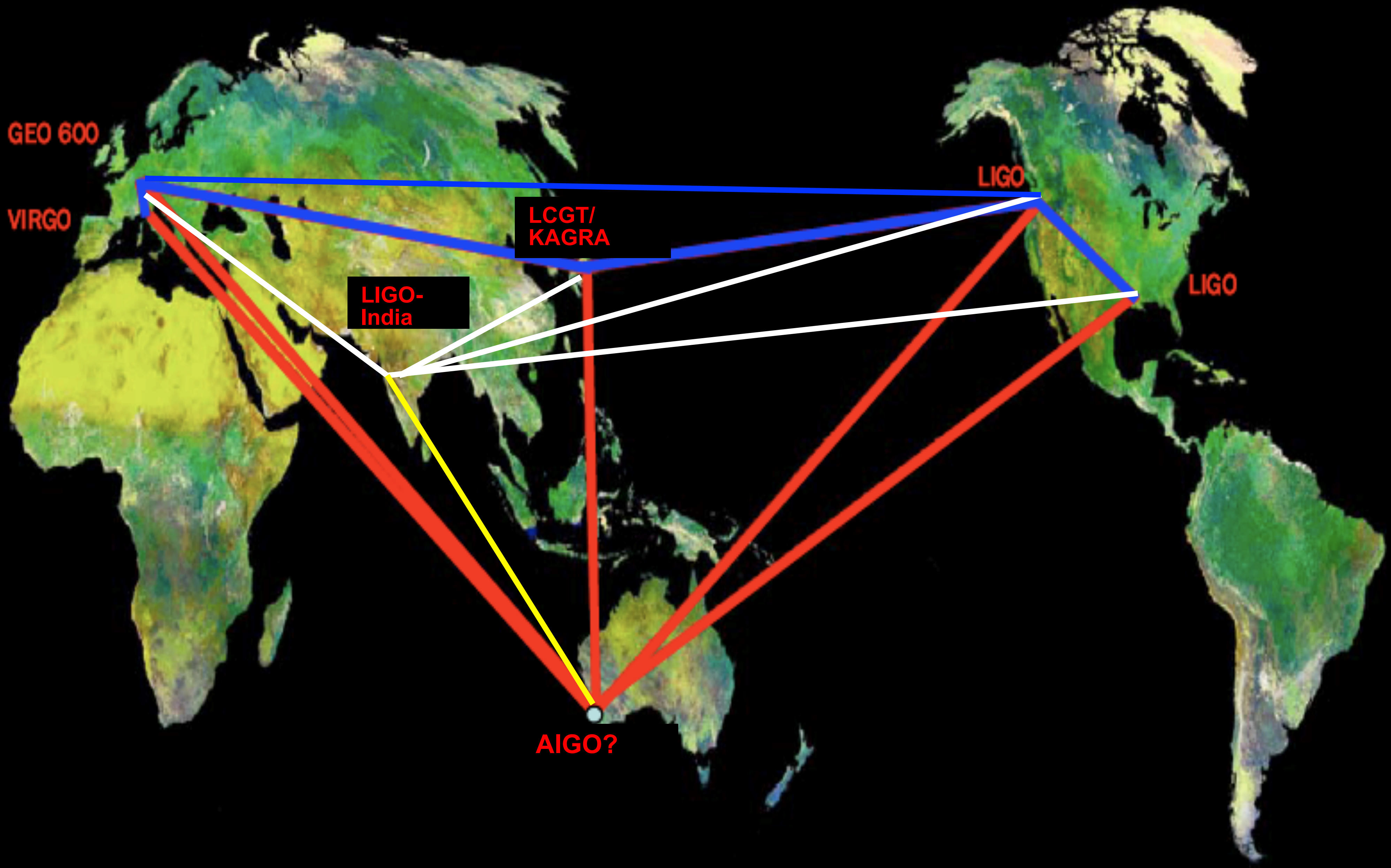
4 km

2 km

4 km

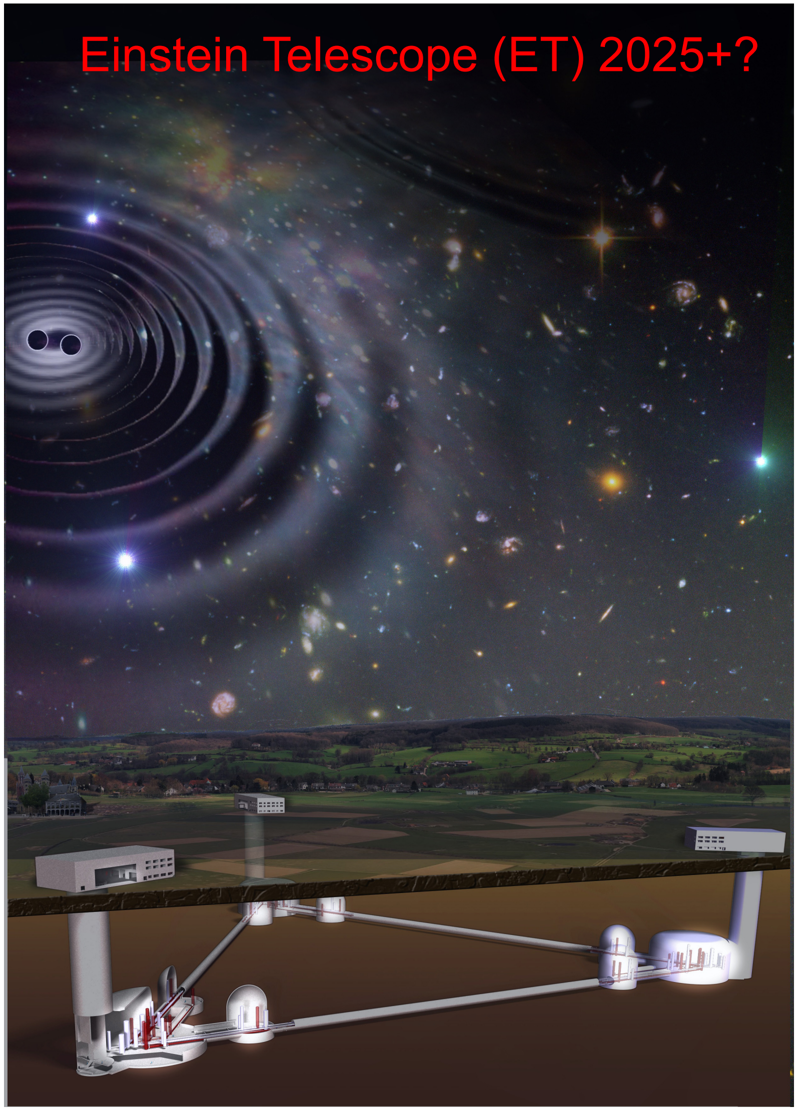
LIGO: Design sensitivity reached ! One year triple coincident data accumulated in 2007

The Global Interferometric GW Detector Network



Long baselines allow source localization using triangulation method

3^{rd} Generation GW Detector



- 10 km long arm
- >150 m underground
- Cryogenic
- Start only after 1^{st} GW detection
- x 10 better sensitivity than aLIGO
- Reach lower frequency of a few Hz
- SKA time scale?

Angular Resolution : Geometrical Expression

- GW with known waveform: angular resolution depends on weighted energy flux coupled to individual detectors
  - 3-site case:

*best at equal detector response

$$\Delta\Omega_s^{(3,best)} = \frac{\pi c^2}{(\xi_1+\xi_2+\xi_3)A_\perp}\sqrt{\frac{1}{\xi_1\xi_2\xi_3/(\xi_1+\xi_2+\xi_3)^3}} \qquad \xi_J \equiv 2\int_{-\infty}^{+\infty}\frac{d\Omega}{2\pi}\Omega^2|\hat{d}_J|^2$$

- GW with unknown waveform: correlation of data or larger number of "null stream" are crucial

$$\Delta\Omega_s^{(\mathrm{Short})} = \frac{4\sqrt{2}\pi c^2}{\sqrt{\sum_{J,K,L,M}\Delta_{JK}\Delta_{LM}|(\mathbf{r}_{KJ}\times\mathbf{r}_{ML})\cdot\mathbf{n}|^2}},$$

$$\begin{aligned}\Delta_{JK} &= \langle\Omega\hat{d}_J|(\delta_{JK}-P_{JK})\Omega\hat{d}_K\rangle \\ &= (\delta_{JK}-P_{JK})\left(2\int_{-\infty}^{+\infty}\frac{d\Omega}{2\pi}\Omega^2\hat{d}_J^*\hat{d}_K\right)\end{aligned}$$

* Also best at equal detector response

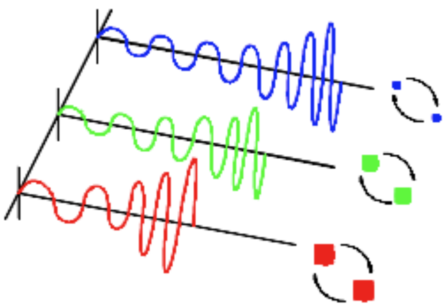
(Wen, Fan & Chen 2008; Wen & Chen 2010)

Ground-based Interferometric GW Detector

- Initial GW detectors are successful
  - Operational at design sensitivity
  - 40 – 2000 Hz
  - Three-detector coincident observation for 1 yr
- Advanced LIGO/VIRGO detectors under construction
  - Operational around 2015
  - 10 – 2000 Hz
  - x 10 improvement in sensitivity, x 1000 in detection volume
  - Tens of detections per year expected

Astrophysical Sources
(Ground-based Detectors)

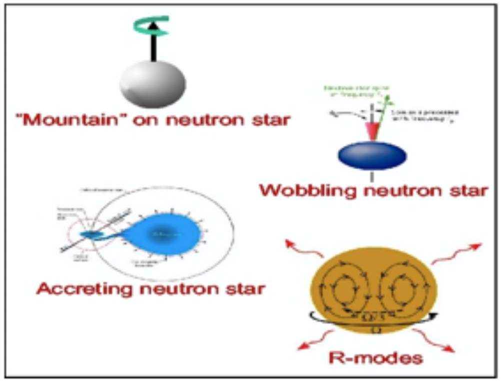
- Compact binary inspiral: *"chirps"*
  - » *NS-NS: waveform known*
    - *same as short gamma-ray bursts?*
    - *expecting optical/radio/X-ray afterglows*
  - » *BH-BH*



- Supernovae / GRBs: *"bursts"*
  - » *not well understood*
  - » *possible*

    E-M counterparts

- Pulsars in our galaxy: *"continuous wave"*
  - » *e.g., mountains on NS*



- Cosmological Signal *"stochastic background"*
  - » *combined GWs from other sources could produce a background of GWs.*
  - » *from the big bang: depend on cosmological models for early universe*

LIGO Data Analysis

- Moderate data rate
  - 16 kHz sampling rate
  - Almost continuous: detector (LV): 70% duty cycle
- Signal is weak
  - Use mostly Maximum Likelihood Ratio principle
  - Suboptimal method to reduce computational cost
- Vetos using environmental channels
- Network analysis a must
    - Coincidence/coherent detection
    - Localization
    - Parameter estimation
    - Veto/Consistency check

Response of the detector to a GW

- Ground-based detectors
  - $d = f^{+} h_{+} + f^{x} h_{x} + n$
    - $f^{+,x}$ is relative orientation between GW and wave frame
      - Detector information known
    - n ~ Gaussian noise

- Pulsar timing array
  - $\Delta\tau = F^{+} g_{+} + F^{x} g_{x} + n_{a} + n$
  - $F^{+,x}$: relative orientation of pulsar-Earth-GW
    - Depending on distance or the pulsar term
  - n_a= "Detector noise"
    - multi-astrophysical effects to be determined

Maximum Likelihood Ratio (MLR) Principle

- Data: x (t) = h (t) +n (t)
- Maximize the Likelihood Ratio

$$\Lambda = P(x \mid h) / P(x \mid 0)$$

- For stationary Gaussian noise

$$\log \Lambda = (x \mid h) - \tfrac{1}{2} (h|h)$$

x: data

s: signal template

(x| y) = correlate x and y weighted by noise

$$\mathrm{E}\left[\tilde{n}(f)\tilde{n}^*(f')\right] = \frac{1}{2} S_n(|f|)\delta(f-f'), \quad -\infty < f, f' < \infty.$$

$$(x|y) := 4\Re \int_0^\infty \frac{\tilde{x}(f)\tilde{y}^*(f)}{S_n(f)}\,\mathrm{d}f.$$

Maximum Likelihood Ratio Principle

- "extrinsic" parameters can be marginalized analytically
- For known waveform
  - If $\mathbf{h}(t;\theta) = a^T h(t;\xi)$
  - solve $\partial \log \Lambda[x; a, \xi]/\partial a = 0$
  - $\log \Lambda = N[x;\xi]^T \, M(\xi)^{-1} \, N[x;\xi]$

    With $N_k[x;\xi] = (x|\mathbf{h}_k(t;\xi))$, $M_{kl}(\xi) = (\mathbf{h}_k(t;\xi)|\mathbf{h}_l(t;\xi))$
- For unknown waveform, marginalize over two wave polarizations
  - burst: $a = (h_+, h_x)^T$ -> excess power method

Data Analysis Method

- Inspiral signal
  - Short duration, antenna beam patters treated as constant
  - Evenly sampled
  - Matched filtering: correlated data with template weighted by noise
  - Search through 2 intrinsic parameters:

    $\xi = (m_1, m_2)$

  - Sub-optimal time-domain IIR filtering method for lowest latency

Challenge/Solution

- Matched filtering method
  - Waveform calculation
    - e.g., BH-BH inspiral/merger waveform
  - Computational issue - Large parameter space
    - Hundreds thousands of templates to be searched for
      - Reduce number of templates
    - Sub-optimal method
    - Bigger computer clusters
    - GPU acceleration
  - Latency issue
    - Overlapping FFT
    - Time-domain method
    - GPU accelerated time-domain search

Time-domain Method
: Summed Parallel IIR Filters

$$y_k = a_k\, y_{k-1} + b_k\, x_k$$

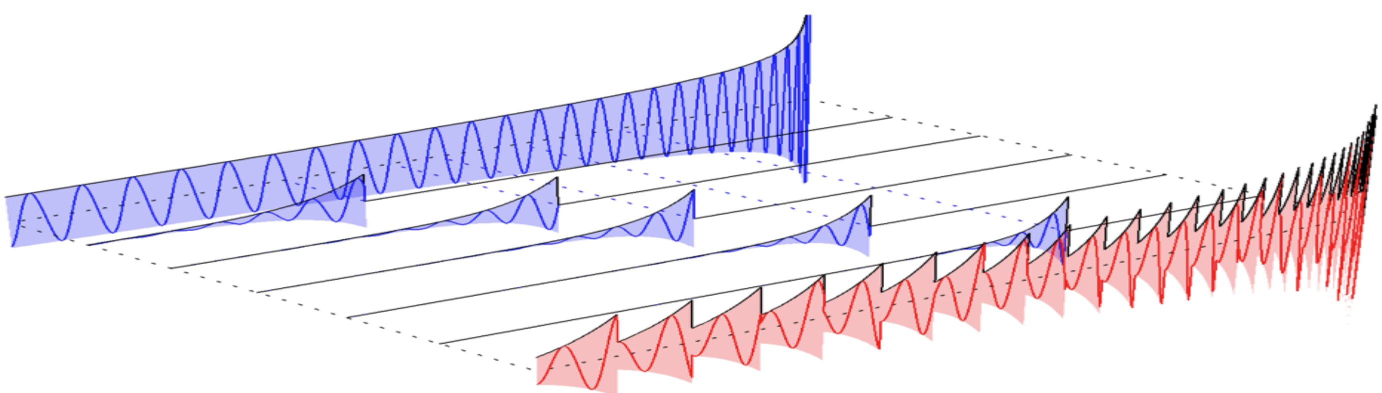
SNR at k-th time step

SNR at (k-1)-th step

k-th GW data

Coefficients calculated for each template at each segment

- Equivalent to matched filter data with a constant-f sinusoid of exponentially rising amplitude (+cutoff)



Luan, J. et al, PRD 2012
Hooper, S. et al, PRD 2012

Data Analysis Method

- Continuous periodic source
  - continuous wave: 4-dim h$(t; \xi)$
    - Long wavelength approximation

$$\hbar 1(t; \xi) = u(t; \xi) \cos \varphi(t; \xi),$$

$$\hbar 2(t; \xi) = v(t; \xi) \cos \varphi(t; \xi),$$

$$\hbar 3(t; \xi) = u(t; \xi) \sin \varphi(t; \xi),$$

$$\hbar 4(t; \xi) = v(t; \xi) \sin \varphi(t; \xi),$$

    - $u(t; \xi)$, $v(t; \xi)$: slowly varying amplitude term
    - $\varphi(t; \xi)$: phase modulation of the signal
    - F-statistics (Stas' talk)

  - Sub-optimal method for all sky search: Hough, Powerflux, Stackslide…
  - Set best spin-down limit for a few sources

Data Analysis Method

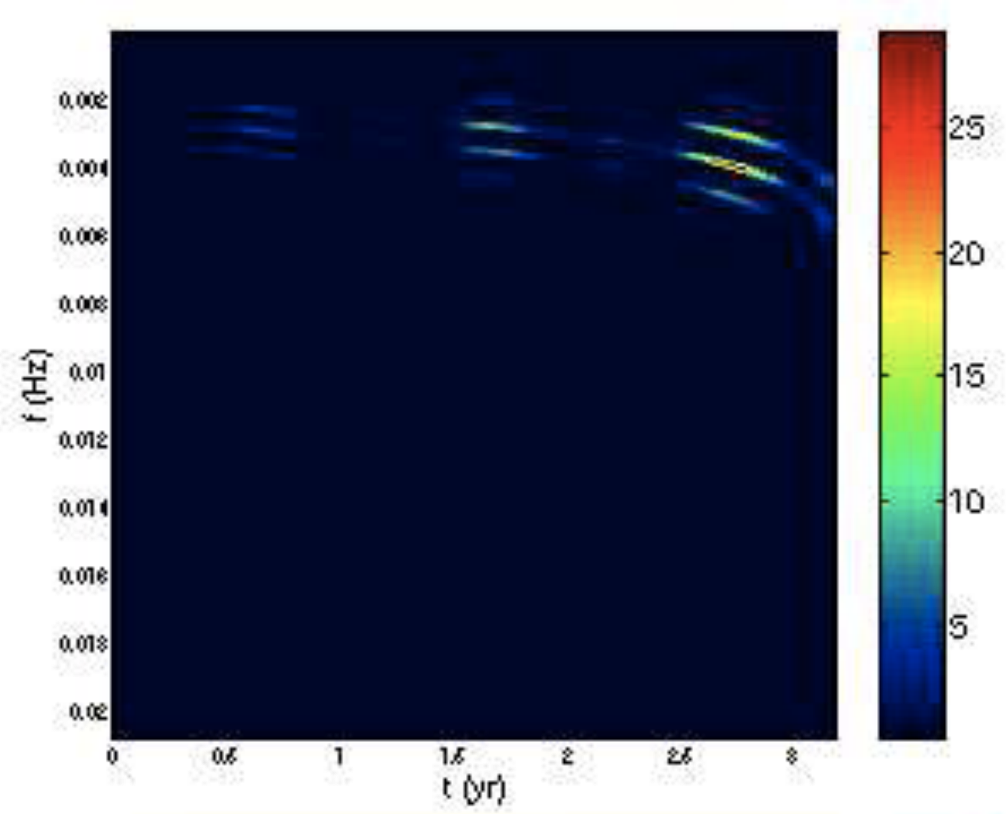
- Stochastic background
  - Correlation method optimized for given model of isotropic GW background

$$\hat{Y} = \int_{-\infty}^{+\infty} df \int_{-\infty}^{+\infty} df'\, \delta_T(f - f')\, \tilde{s}_i^\star(f)\, \tilde{s}_j(f')\, \tilde{Q}_{ij}(f')$$

    - e.g., Marginalized over calibration error
    - Better limit than indirect Big Bang nuclear synthesis and cosmic microwave background at 100 Hz
  - Radiometer
    - Directional correlation
    - For not-so-isotropic background

Data Analysis Method

- Burst source
    - Excess power
    - Time-frequency method or the kind



Network Analysis

Network Analysis

- Maximum Likelihood Ratio
  - Noise in different detectors are assumed to be uncorrelated
  - For each sky direction, wave-front arrival time can be calculated between detectors
  - Data d are time-shifted for each sky direction

$$\hat{\mathbf{d}} = \hat{A}\mathbf{h} + \hat{\mathbf{n}},$$

$$\mathbf{A} = (\mathbf{f}^{+}, \mathbf{f}^{x})$$

$$\mathbf{h} = (\mathbf{h}_{+}, \mathbf{h}_{x})$$

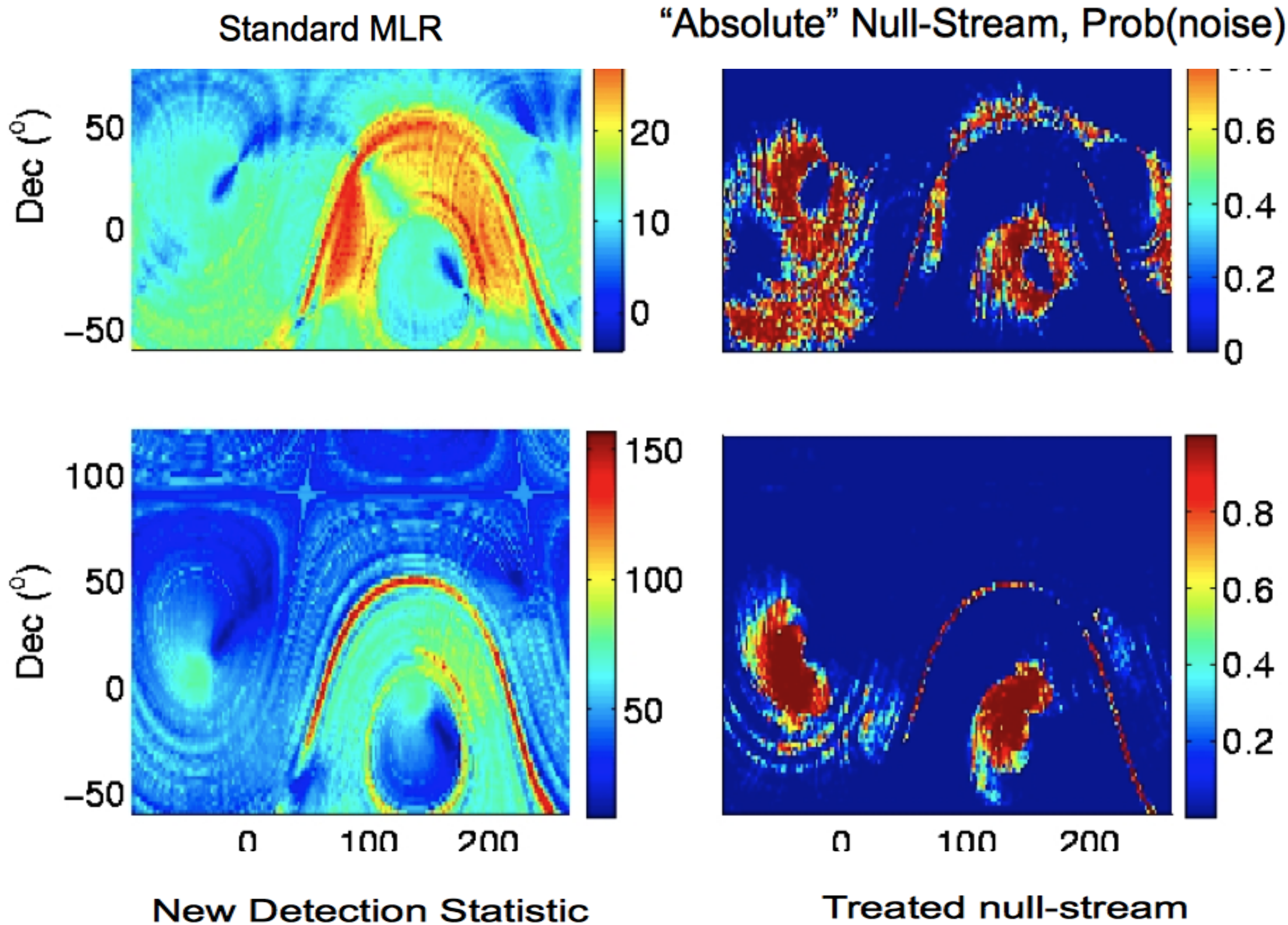
Network Analysis

- Data from different detectors can be linearly recombined to form signal streams and null streams

$$\hat{\mathbf{d}}' = u^\dagger \hat{\mathbf{d}}, \quad \mathbf{h}' = v^\dagger \mathbf{h}, \quad \hat{\mathbf{n}}' = u^\dagger \hat{\mathbf{n}}$$

$$\begin{pmatrix} \hat{d}'_1 \\ \hat{d}'_2 \\ \dots \\ \dots \\ \dots \\ \dots \\ \hat{d}'_{N_k N_d} \end{pmatrix} = \begin{pmatrix} s_1 & 0 & \dots & 0 \\ 0 & s_2 & \dots & 0 \\ 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & s_{2N_k} \\ 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 \end{pmatrix} \begin{pmatrix} h'_1 \\ h'_2 \\ \dots \\ h'_{2N_k} \end{pmatrix} + \begin{pmatrix} \hat{n}'_1 \\ \hat{n}'_2 \\ \dots \\ \dots \\ \dots \\ \dots \\ \hat{n}'_{N_k N_d} \end{pmatrix}$$

Wen & Schutz (2005), Wen & Schutz (2012)



LHG, SNR ~ 20, SNR(GEO)~2, src location in center

PTA

- Maximum Likelihood Ratio
  - Seen in many publications already

    e.g., van Haasteren et al 2011, Babak, et al., siemens et al. see Justin's talk
  - Suboptimal methods

- Network Analysis
  - For each pulsar, data from different detectors added (Monday's talk)
  - For each source direction, can data from different pulsars be added coherently?
  - Distance to pulsar is the key
  - Geometrical expression for angular resolution can be obtained

Conclusion

- LIGO and PTA response to GWs are similar
- For PTA:
  - Need to detect the “detector” behavior
  - distance to pulsars is the key
    - Network analysis
    - Angular resolution
- Maximum Likelihood Ratio methods are for sensitive search of weak signals
  - Analytical maximization helps
- Sub-optimal methods also used for different purposes
- Computational efficiency can be addressed
  - More computer clusters
  - GPU acceleration
  - Smarter way of laying templates, etc