Some ideas about how to compare distributions

The Niels Bohr International Academy

NATIONAL SECURITY AGENCY CENTRAL SECURITY SERVICE

Defending Our Nation. Securing The Future.

DR. SOLOMON KULLBACK

1907-1994

DR. RICHARD A. LEIBLER

1914-2003

The Niels Bohr International Acad

Kullback-Leibler divergence (1951):

$$D_{\mathrm{KL}}(P\|Q) = \sum_i P(i)\,\ln\frac{P(i)}{Q(i)} \qquad \text{or :} \quad \int_{-\infty}^{\infty} p(x)\,\ln\frac{p(x)}{q(x)}\,\mathrm{d}x$$

entropy

cross entropy

Lots of nice properties:

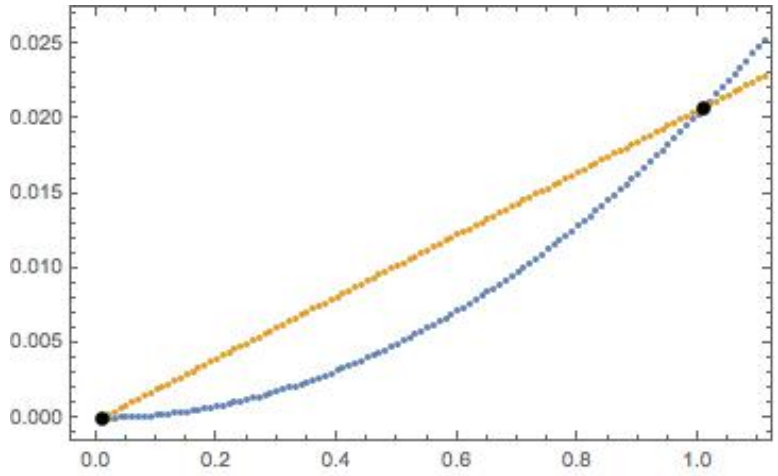
$$D_{\mathrm{KL}}(P\|Q) \geq 0$$ since ln (x) ≤ x - 1

- Invariance under parameter changes:

$$\int_{x_a}^{x_b} P(x)\log\left(\frac{P(x)}{Q(x)}\right)dx = \int_{y_a}^{y_b} P(y)\log\left(\frac{P(y)dy/dx}{Q(y)dy/dx}\right)dy = \int_{y_a}^{y_b} P(y)\log\left(\frac{P(y)}{Q(y)}\right)dy$$

The KL divergence is convex:

Consider D_{KL}(p, [1-x]p+xq) for 0 < x < 1. Convexity says that this divergence is always less than the linear interpolation for any choice of the distributions p and q.



Convexity ensures that Jensen's inequality, Hölder's inequality, and the arithmetic-geometric mean inequality all hold. Check this out on the Wikipedia.

arith.-geom. $$\frac{1}{N}\sum_{i=1}^{N} x_i \geq \left(\prod_{i=1}^{N} x_i\right)$$

$$\mathbf{E}[x] = \sum_i x_i p_i$$

$$\mathbf{E}[|xy|] \leq \left(\mathbf{E}[|x^r|]\right)^{1/r} \left(\mathbf{E}[|y^s|]\right)^{1/s}$$ Hölder

Some ideas about
how to compare distributions

The Niels Bohr
International Academy

Some ideas about
how to compare distributions

The Niels Bohr
International Academy

- It is additive for independent distributions:

If $P(x,y) = P_1(x)P_2(y)$ and $Q(x,y)=Q_1(x)Q_2(y)$,

$$D_{\mathrm{KL}}(P\|Q) = D_{\mathrm{KL}}(P_1\|Q_1) + D_{\mathrm{KL}}(P_2\|Q_2).$$

The KL divergence is *not* a metric since $D_{\mathrm{KL}}(P\|Q) \neq D_{\mathrm{KL}}(Q\|P)$

...but symmetry is restored when $Q(i) = P(i) + \delta P(i)$ and $\delta P(i) \to 0$,

$$D_{KL}(P \parallel Q) \to \frac{1}{2} \sum_i \frac{\delta P(i)^2}{P(i)}$$

For N independent draws we expect that $n_i = NP(i) \pm \sqrt{NP(i)}$ so that the proxy for $P(i)$ is $P(i) \pm \sqrt{P(i)/N}$. All bins are created equal!

- KL divergence also fails to satisfy the triangle inequality

$$D_{KL}(P \parallel Q) \leq D_{KL}(P \parallel R) + D_{KL}(R \parallel Q)$$

- So, the KL divergence is sometimes called a `premetric'.

In spite of its formal shortcoming, the KL divergence can be an extremely useful tool for comparing distributions.

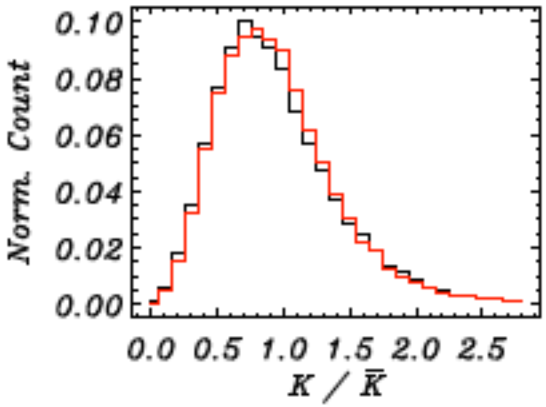
A simple physical interpretation:

For N uncorrelated random draws on a distribution P, the most probable result will be $n_i = NP(i)$. The probability that this "most probable" draw will be drawn on P is Π_P; the probability that it will be drawn on Q is Π_Q. The KL divergence can then be written as

$$D_{KL}(P \parallel Q) = -\frac{1}{N} \ln \left(\Pi_P / \Pi_Q \right) .$$

The Niels Bohr International Academy

Often, we are interested in comparing P with the distribution obtained from result of N random draws on P. How do the statistical properties of D_{KL} depend on N?



N	$\overline{K}$	ΔK	$N\overline{K}$	$N\Delta K$
4000	0.002540	0.000796	10.160	3.18
8000	0.001253	0.000413	10.024	3.30
16000	0.000636	0.000200	10.176	3.20

Mean and RMS values of the KL divergence for a discrete Gaussian distribution.

If the data is sorted into N_{bin} bins, we should expect that D_{KL} will be of order N_{bin}/N.

Vol 444|21/28 December 2006 nature

COMMENTARY

Measures for measures

Are some ways of measuring scientific quality better than others? **Sune Lehmann**, **Andrew D. Jackson** and **Benny E. Lautrup** analyse the reliability of commonly used methods for comparing citation records.

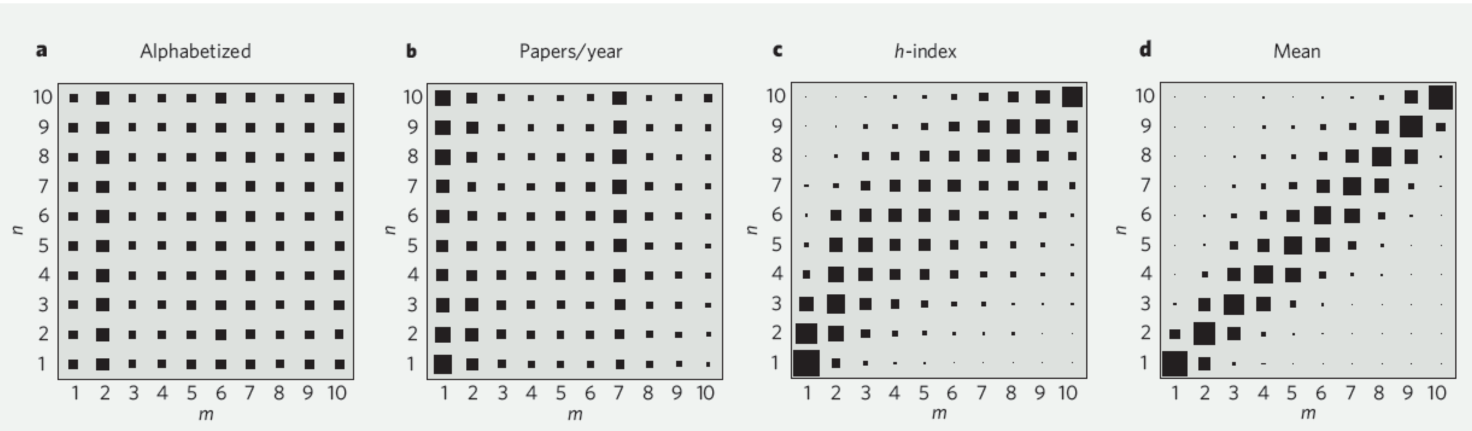


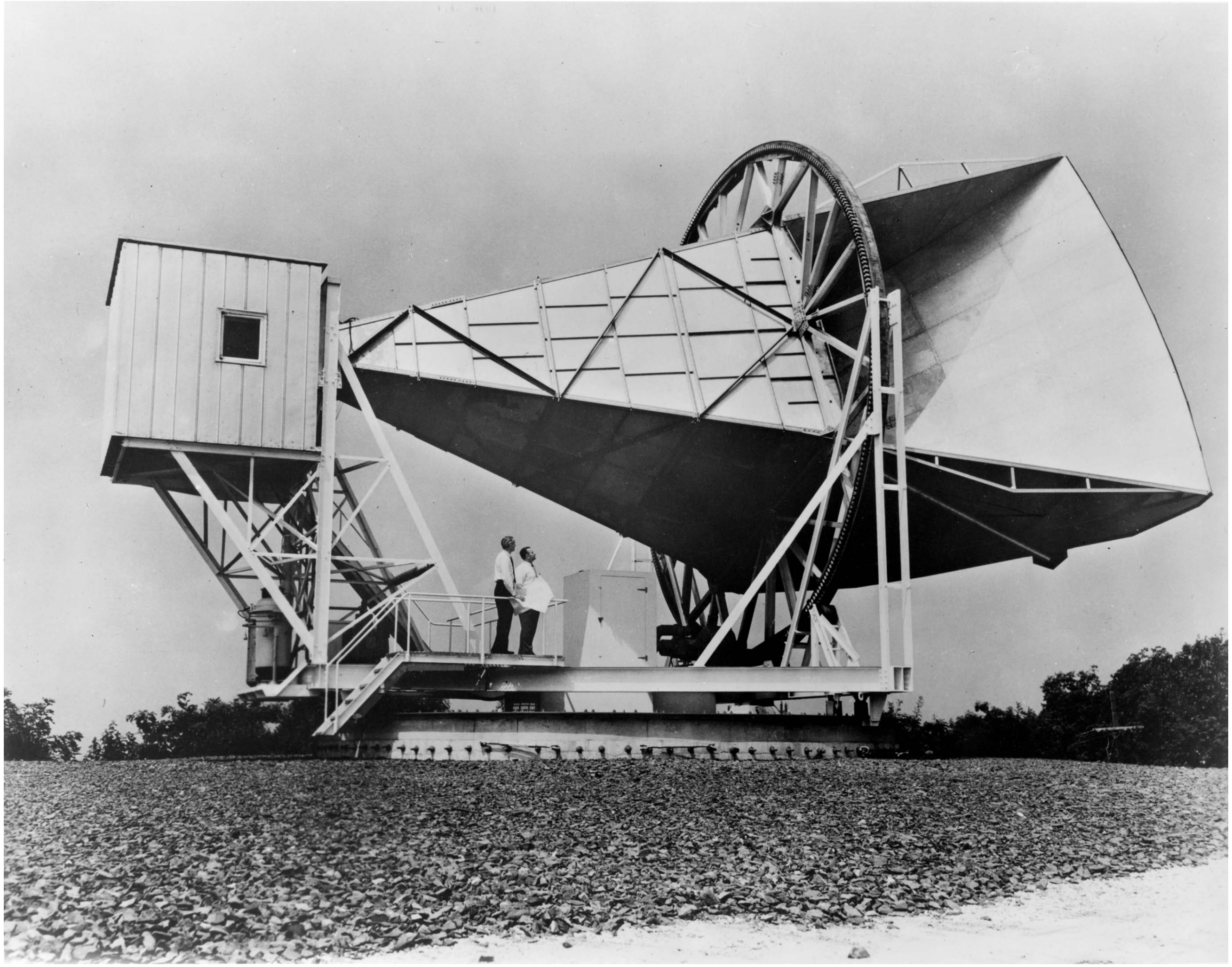
Figure 1 | The probabilities for four different measures. a–d, Each horizontal row, indexed by n, shows the average probabilities that authors initially assigned to a given decile bin n are predicted to lie in a different decile bin m. The probabilities are proportional to the areas of the corresponding squares.

The Niels Bohr International Academy

So, what else can spies do for us?

The Niels Bohr
International Academy

NASA’s Echo I and Echo II “satelloon” program (1960-69)

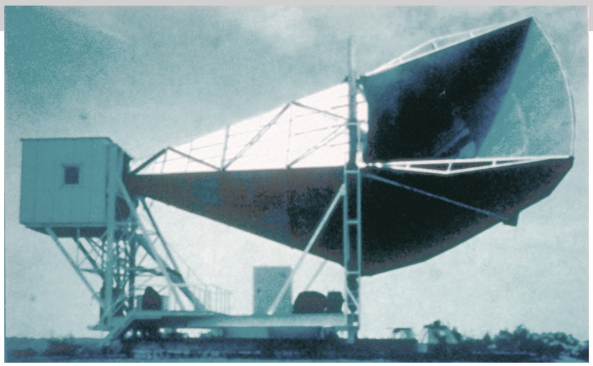


The Holmdel Horn

Arno Penzias and Robert Wilson

Completely uniform microwave background at 300 GHz!

1965 **Penzias and Wilson**



Princeton to the rescue … In early 1960s Dicke, Zeldovich and Igor Novikov independently suggested such radiation as a remnant of the Big Bang some 13.8 billion years ago!

ROBERT HENRY DICKE

6 MAY 1916–4 MARCH 1997

The Niels Bohr International Academy

Penizias and Wilson knew *nothing* about CMB:

free from seasonal variations (July, 1964–April, 1965). A possible explanation for the observed excess noise temperature is the one given by Dicke, Peebles, Roll, and Wilkinson (1965) in a companion letter in this issue.

They didn't know how to say "thank you" either:

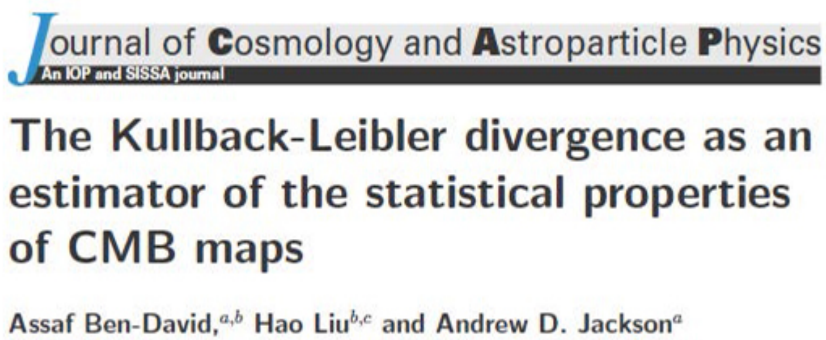
We are grateful to R. H. Dicke and his associates for fruitful discussions of their results prior to publication. We also wish to acknowledge with thanks the useful comments and advice of A. B. Crawford, D. C. Hogg, and E. A. Ohm in connection with the problems associated with this measurement.

ROBERT HENRY DICKE

6 MAY 1916–4 MARCH 1997

The Niels Bohr International Academy

SMICA map of CMB temperature fluctuations:

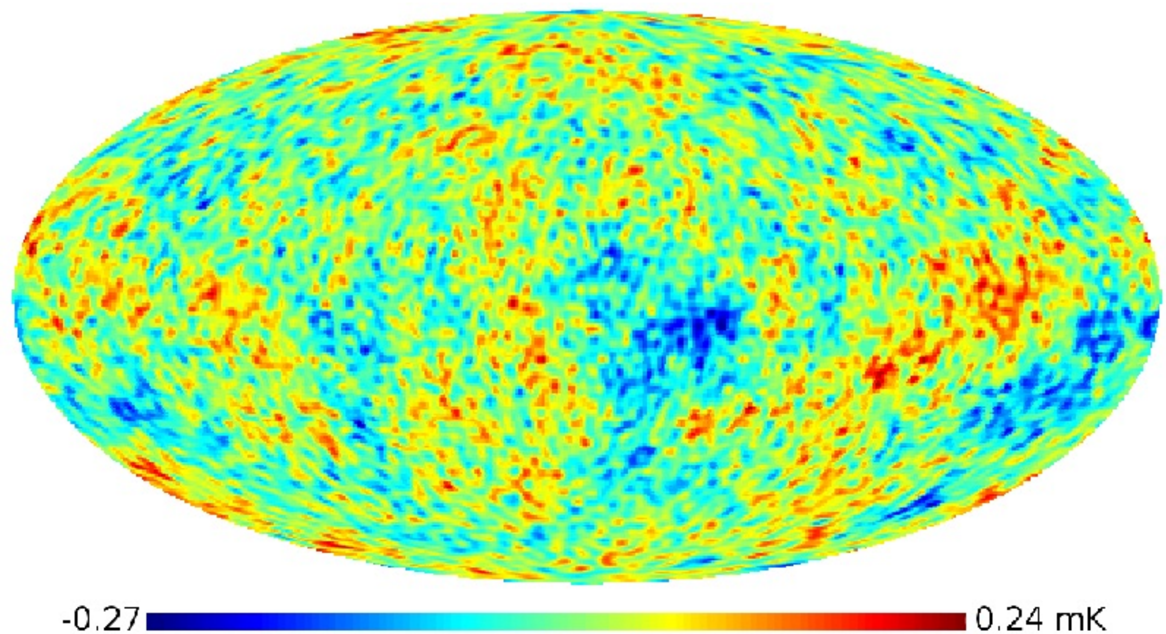

Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

The Kullback-Leibler divergence as an estimator of the statistical properties of CMB maps

Assaf Ben-David,[a,b] Hao Liu[b,c] and Andrew D. Jackson[a]

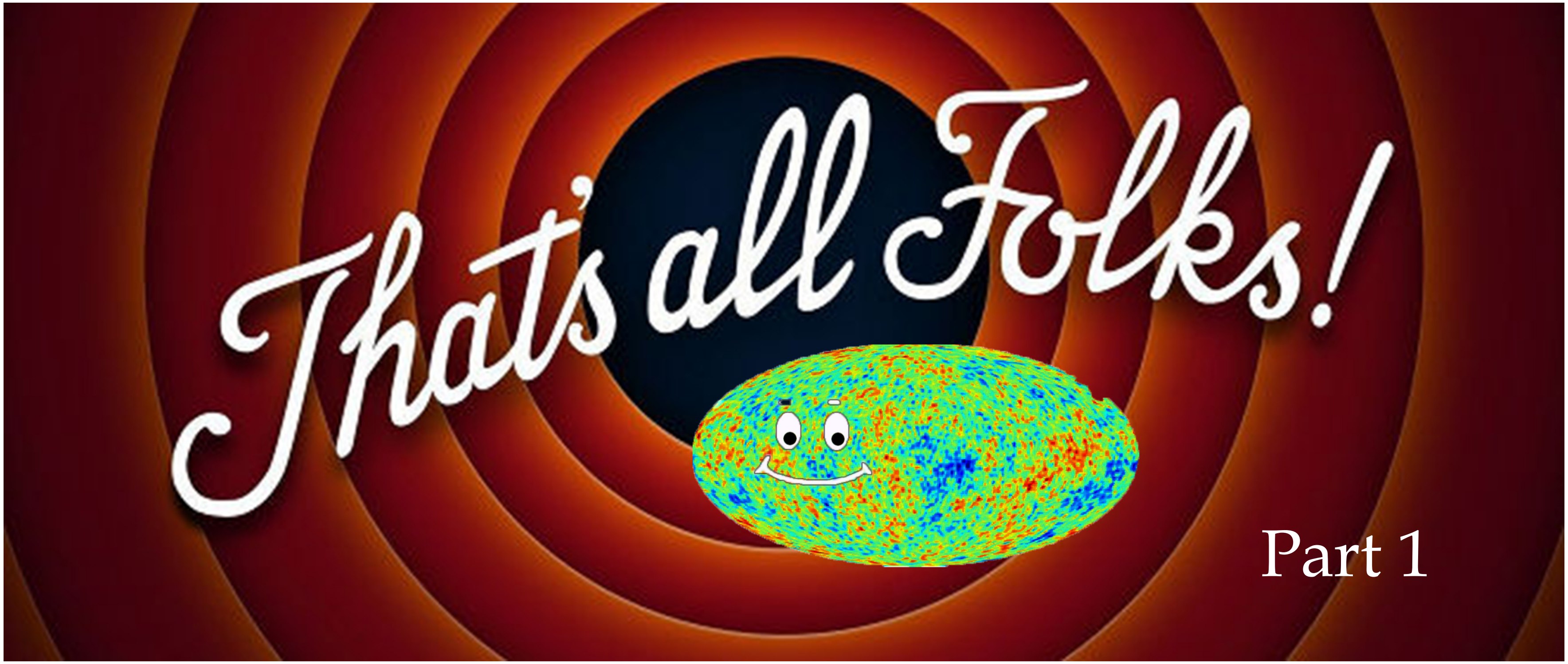
and  = **Pavel Naselsky**

Is the data Gaussian?



The Niels Bohr International Academy

KL says that it is.

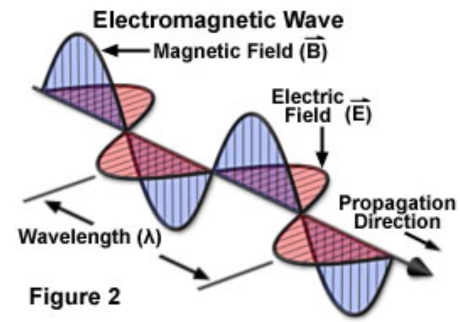


The Niels Bohr International Academy

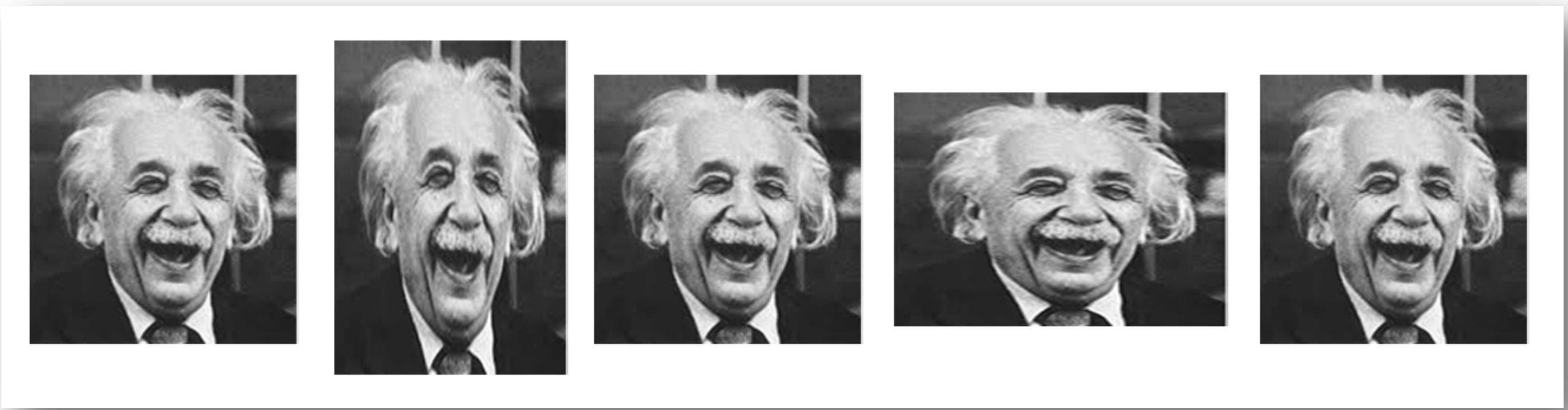
The Niels Bohr International Academy

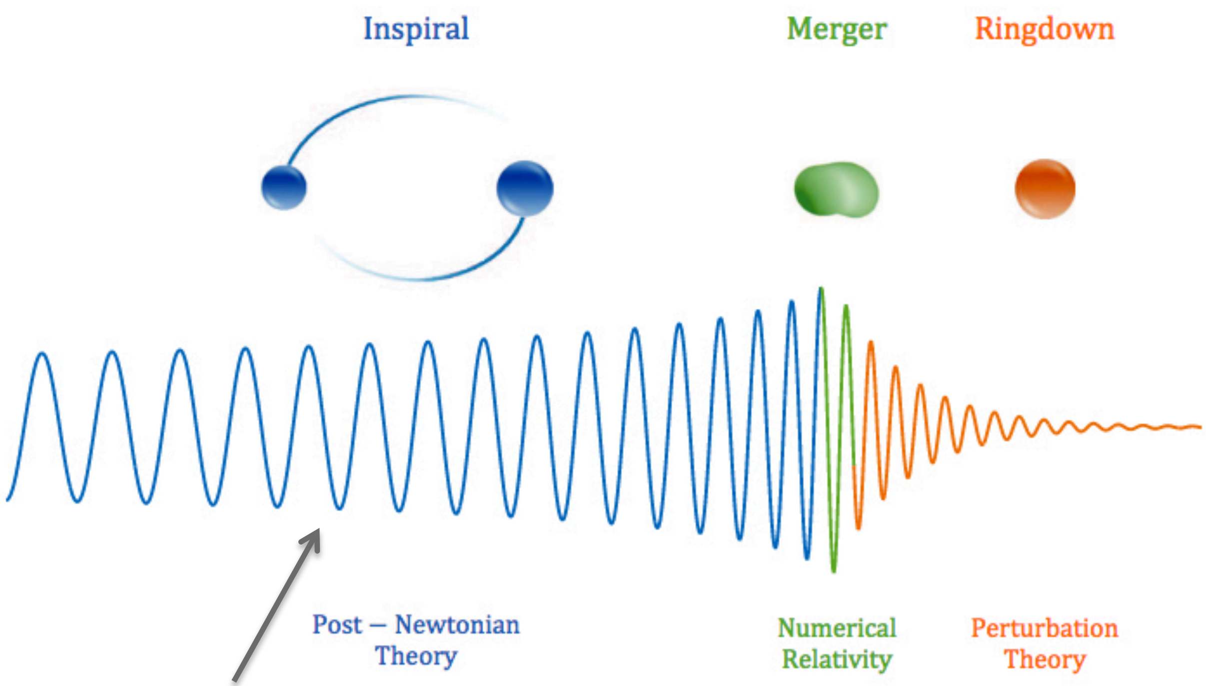
Gravitational Waves:

Accelerating *charges* radiate transverse electromagnetic waves.



Accelerating *masses* radiate gravitational waves that are also transverse. A gravitational wave in direction z will alternately squeeze and stretch space in the perpendicular directions x and y.





$$\left(\frac{5\tau}{t_0 - t}\right)^{1/4} \cos\left[2\left(\frac{t_0 - t}{5\tau}\right)^{5/8} - \phi_0\right]$$

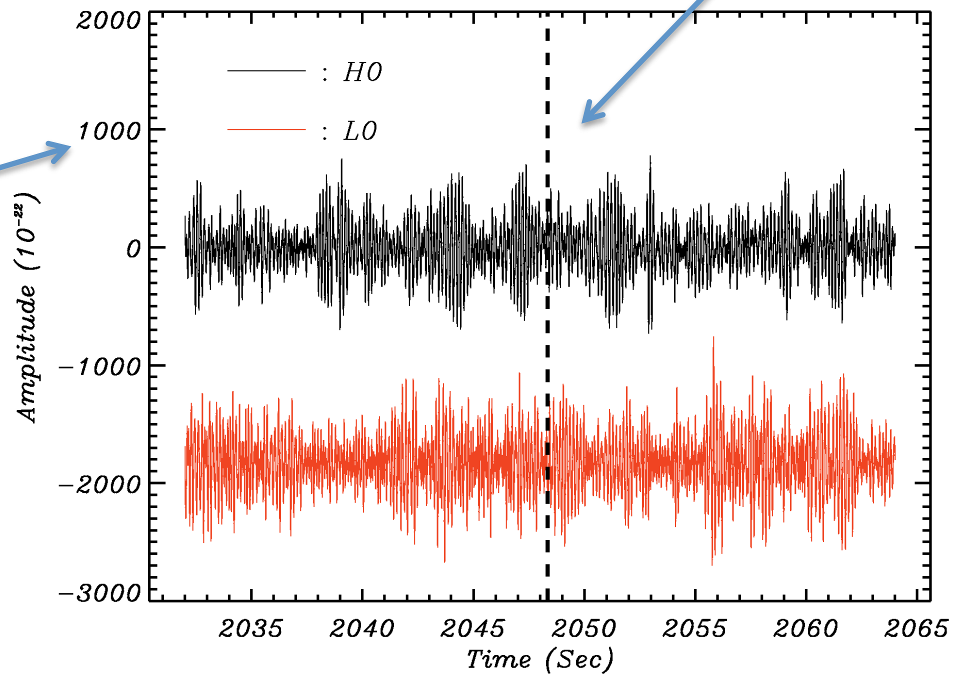
with $\tau = G\mathcal{M}_{\rm ch}/c^3$ and $\mathcal{M}_{\rm ch} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$

(This wave form is characteristic for any “catastrophic” event.)

The initial 32 s time records for Hanford and Livingston

See the event?

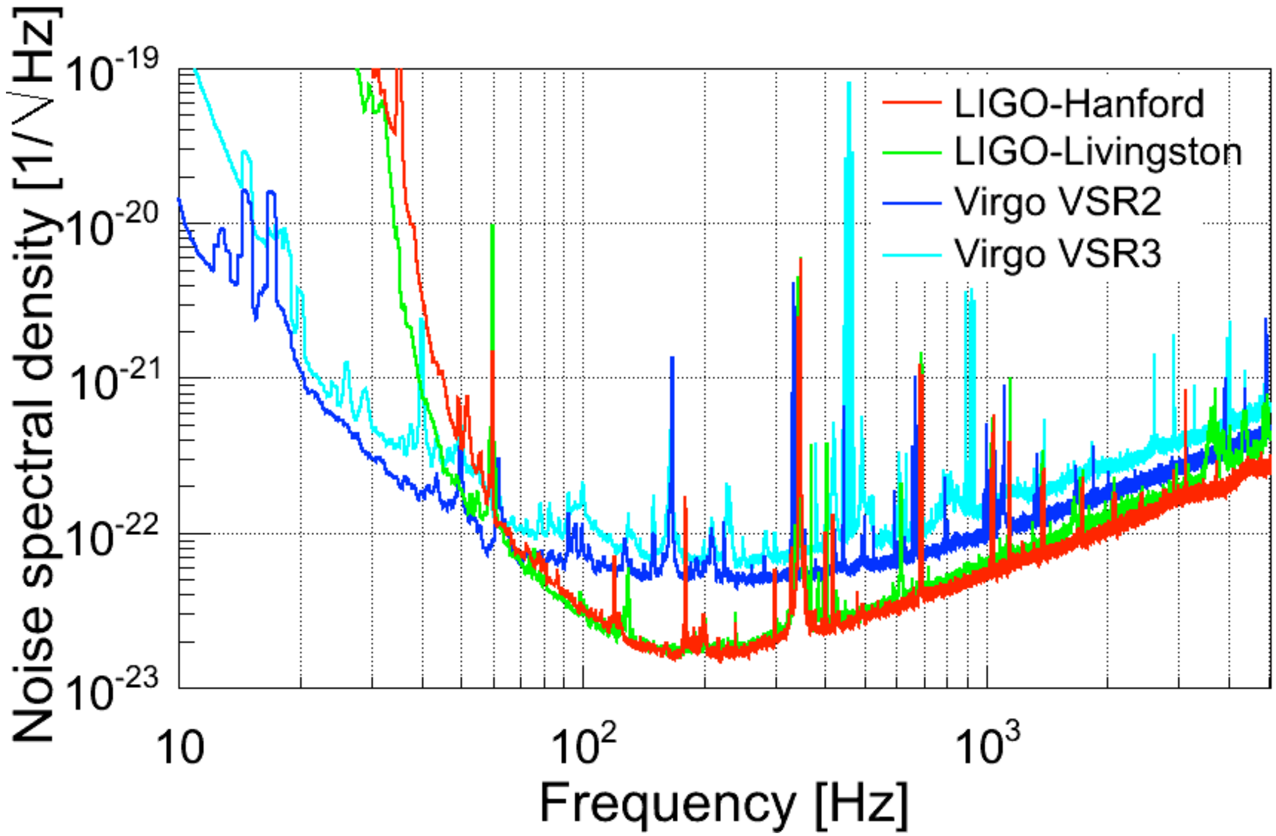
See the scale!



“Noise” is roughly 300 times larger than “signal”. Cleaning is required! Unfortunately, the available data has been “pre-cleaned” with unknown consequences.

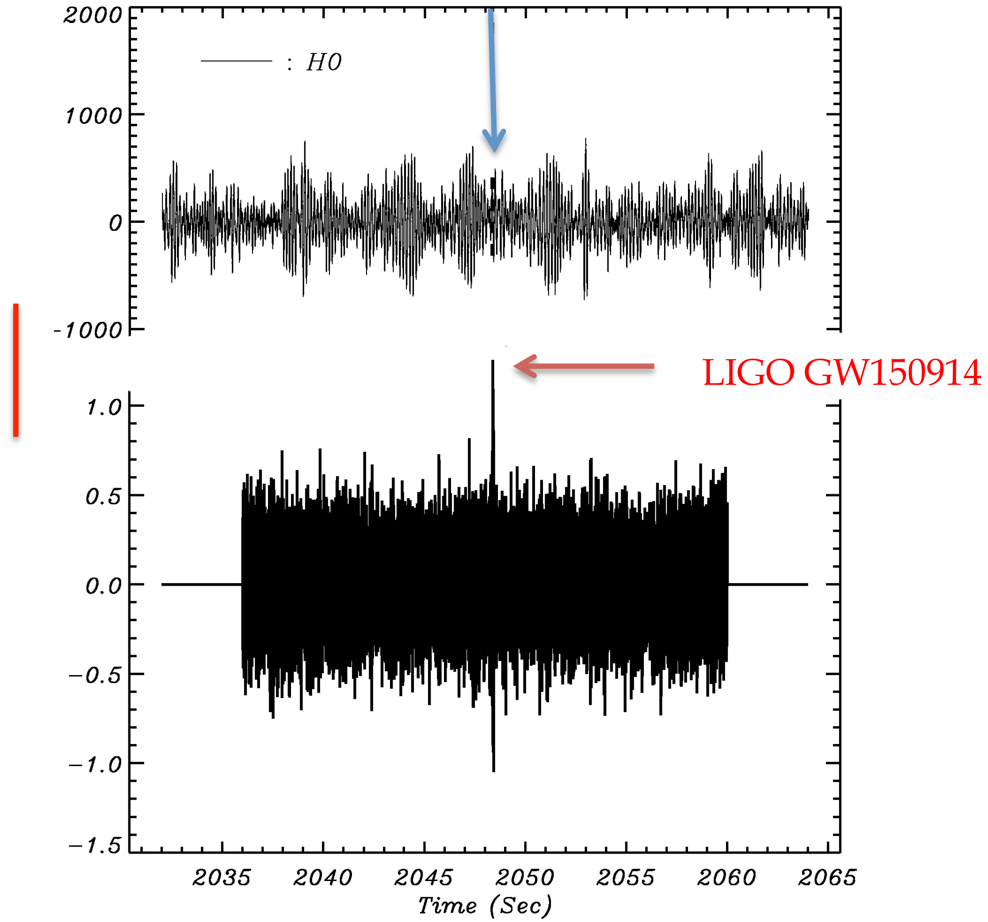
The Niels Bohr International Academy

High sensitivity = Noise! This includes narrow resonances (e.g., calibration signals and 60 Hz noise and harmonics).



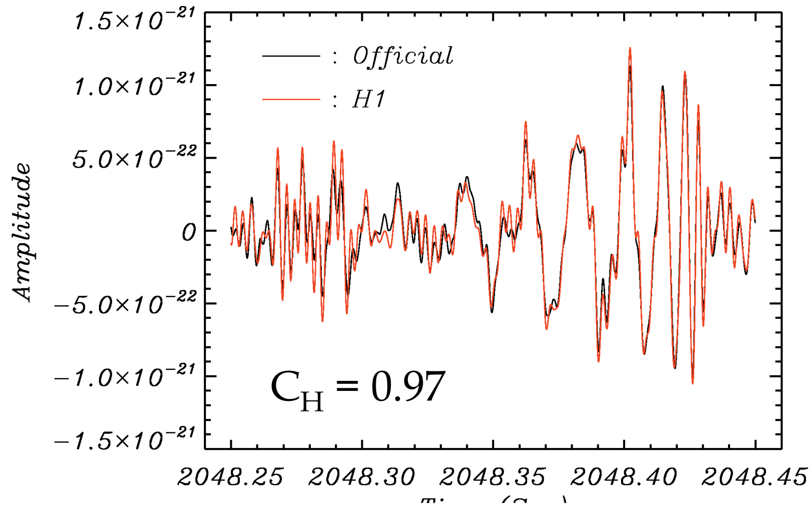
Data cleaning is essential.

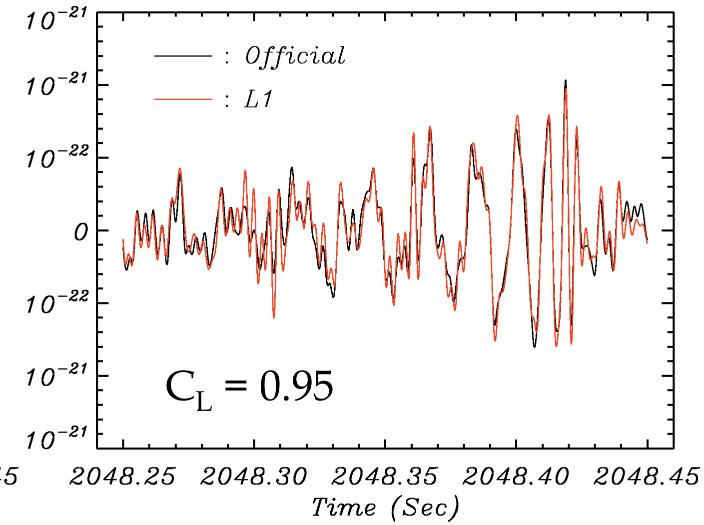
The Niels Bohr International Academy

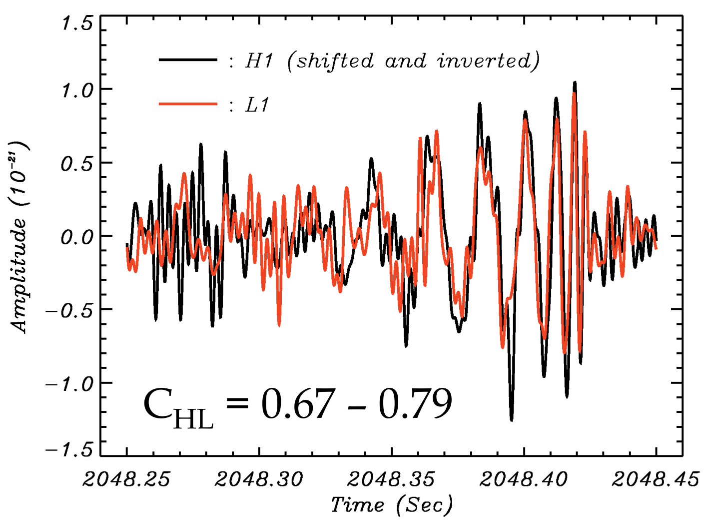


The simplest possible cleaning is sufficient to reveal the signal!

The signals …







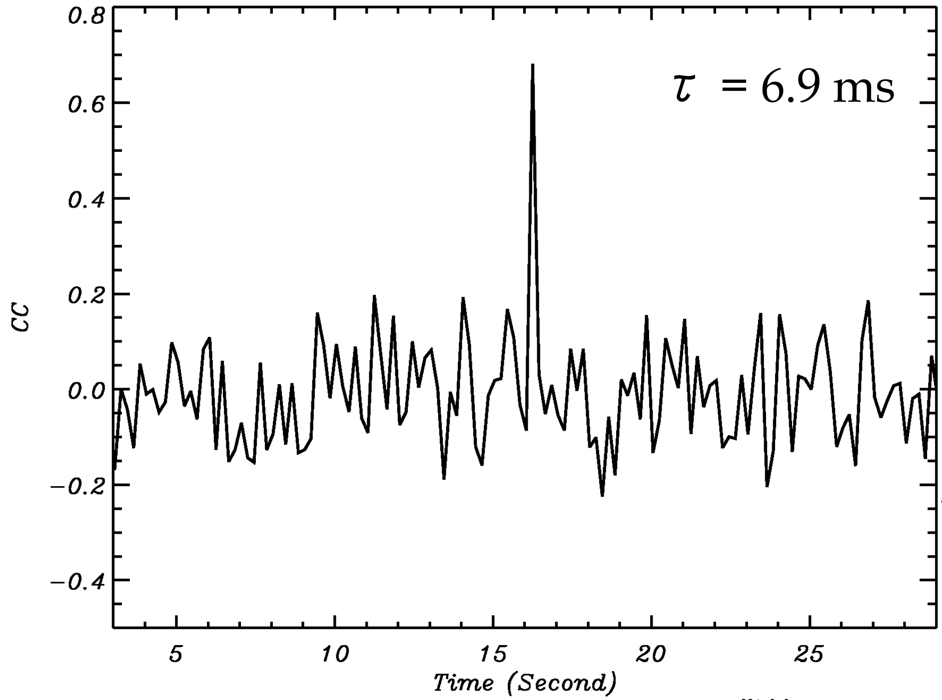
Masses are 36^{+5}_{-4} and $29^{+4}_{-4}\, M_\odot$ with some $3\, M_\odot c^2$ emitted as gravitational radiation.

The Niels Bohr International Academy

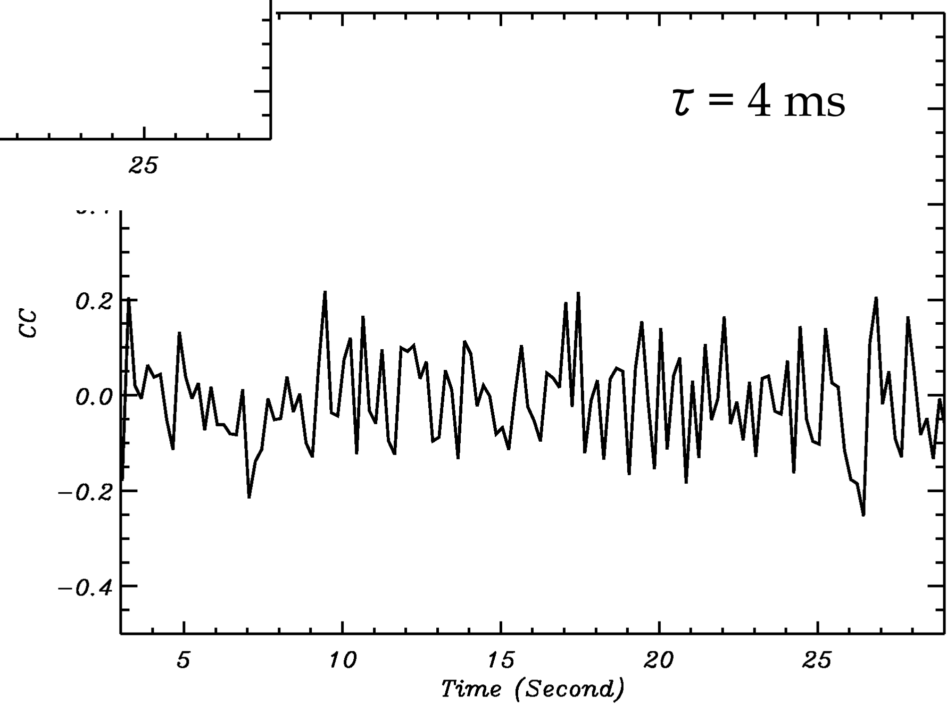
Pearson cross-correlation coefficient:

$$
\begin{aligned}
C(\tau) &= \frac{1}{N} \sum_j \overline{s}_H(t_j)\overline{s}_L(t_j, \tau), \\
N &= \left(\sum_i \overline{s}_H^2(t_i) \; \sum_k \overline{s}_L^2(t_k, \tau) \right)^{1/2} \\
\overline{s}_{H,L} &= \overline{S}_{H,L} - \langle \overline{S}_{H,L} \rangle \, .
\end{aligned}
$$

Here, τ is the time shift between the two detectors.



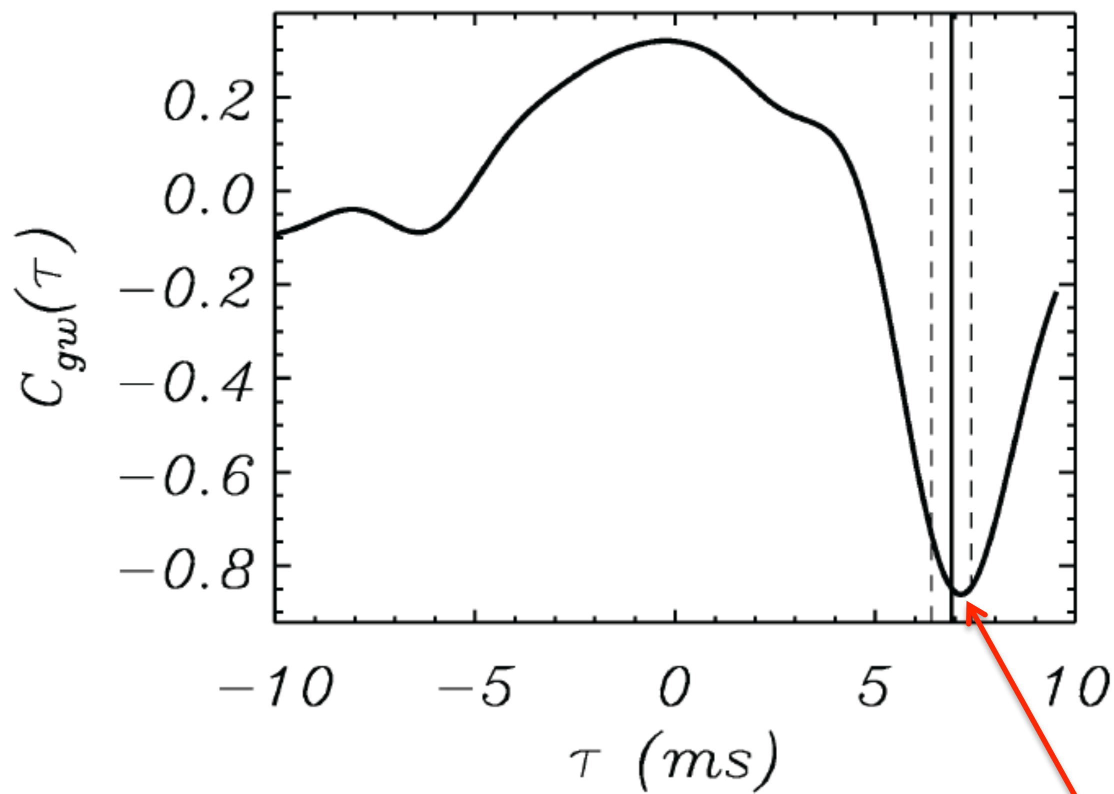
τ = 6.9 ± 0.4 ms



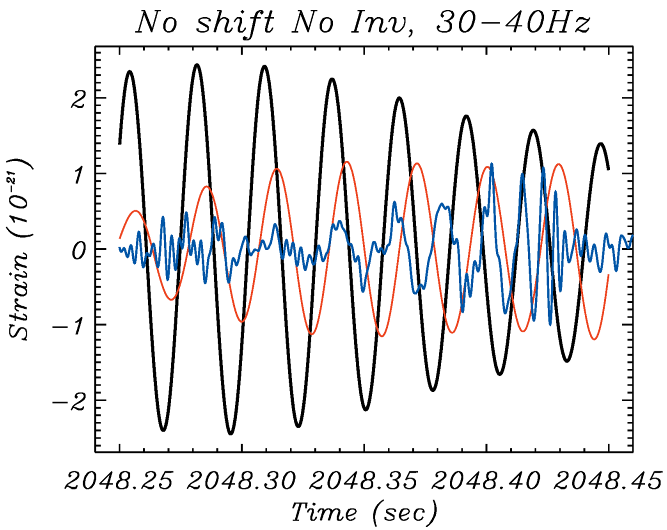
C(τ) makes it is simple to determine the time delay between H and L.

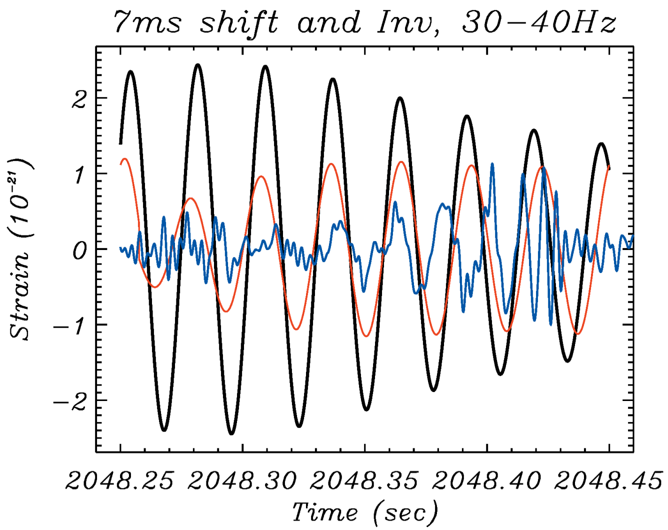
The Niels Bohr International Academy

The cross-correlator between Hanford and Livingston signals as a function of the time delay.



"confirmation" of the event



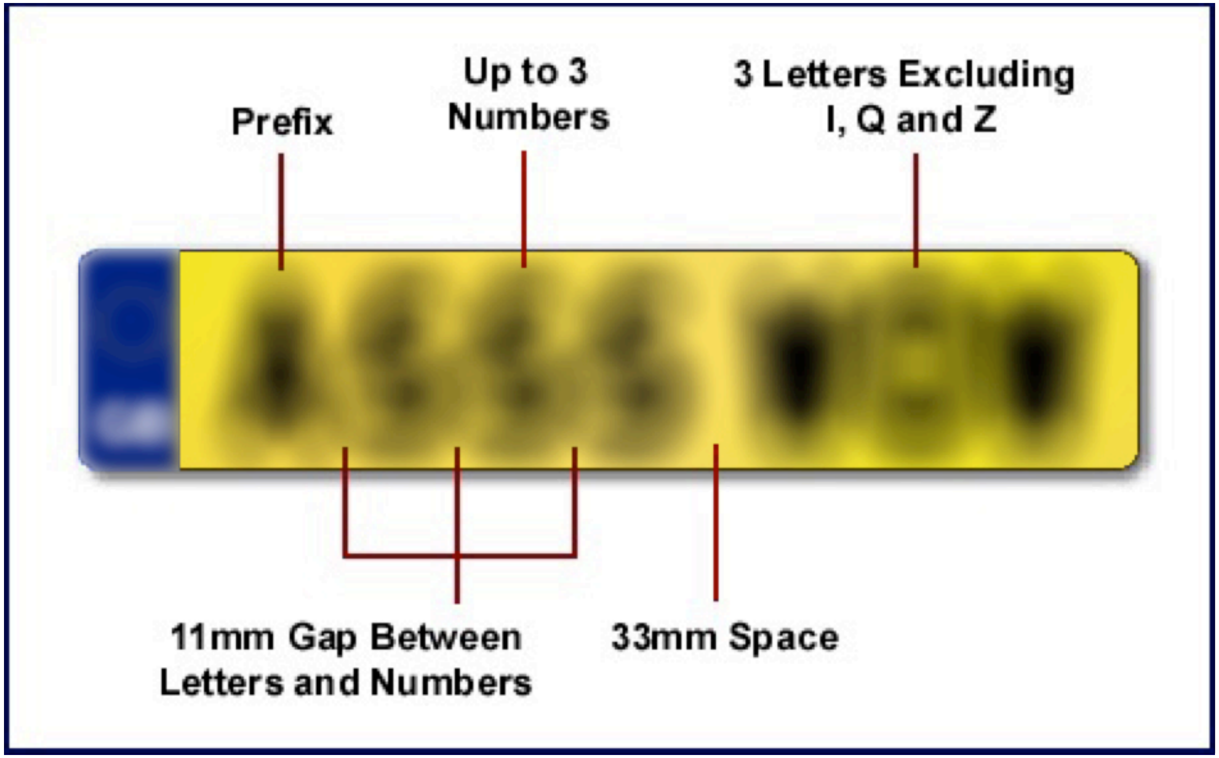


Hanford has calibration lines at 35.9 and 36.7 Hz; Livingston has similar lines at 34.7 and 35.3 Hz.

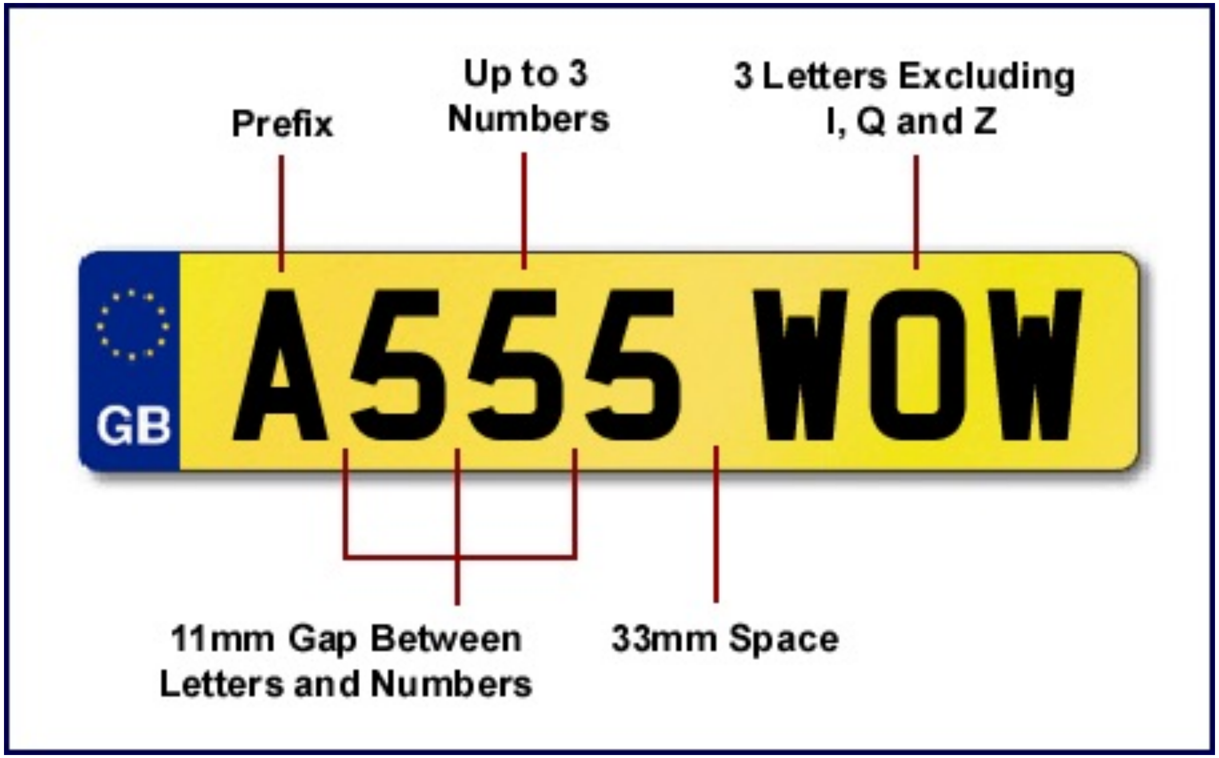
The *same* 7 ms shift brings these lines in phase.

The Niels Bohr International Academy

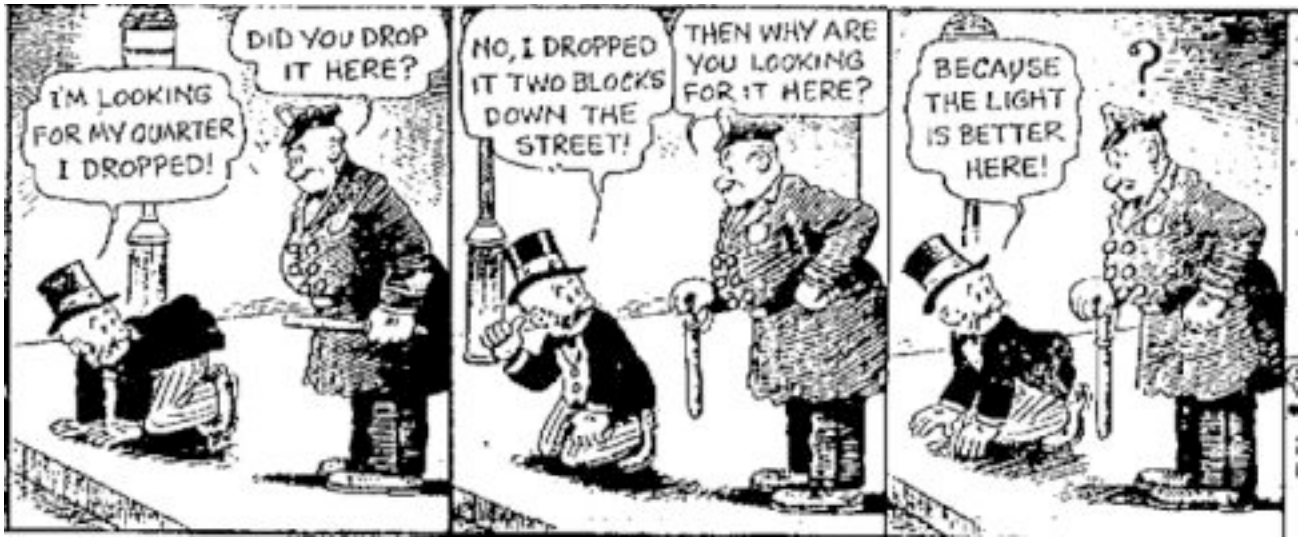
Police routinely use template analysis because they know *exactly* what to look for.



But a French license plate would cause problems.



We must know the probability that an observed event
is a gravitational wave and that it is due to black hole merger.



The Niels Bohr International Academy

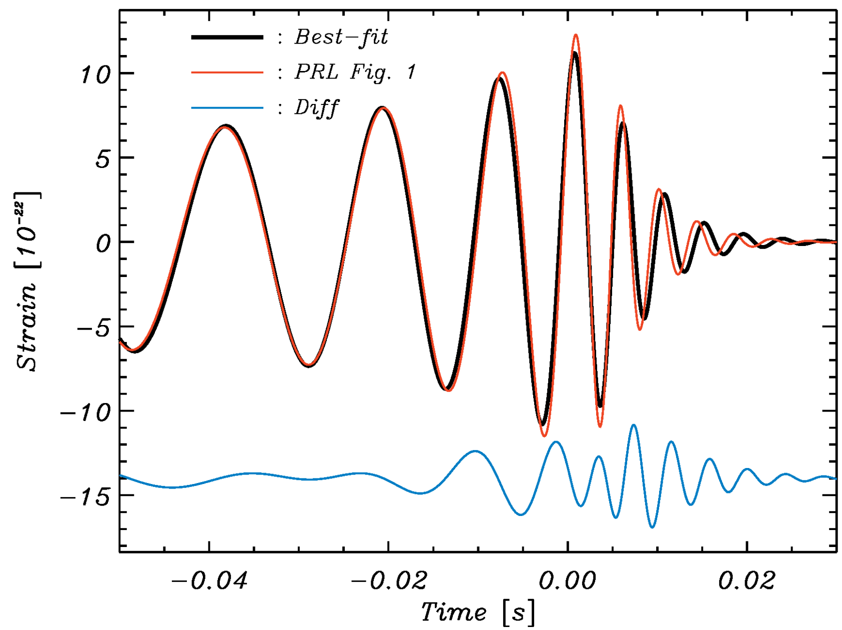
The Secret Template

LOSC_Event_tutorial_fixed.py × LOSC_Event_tutorial.py × whiten_simple.py

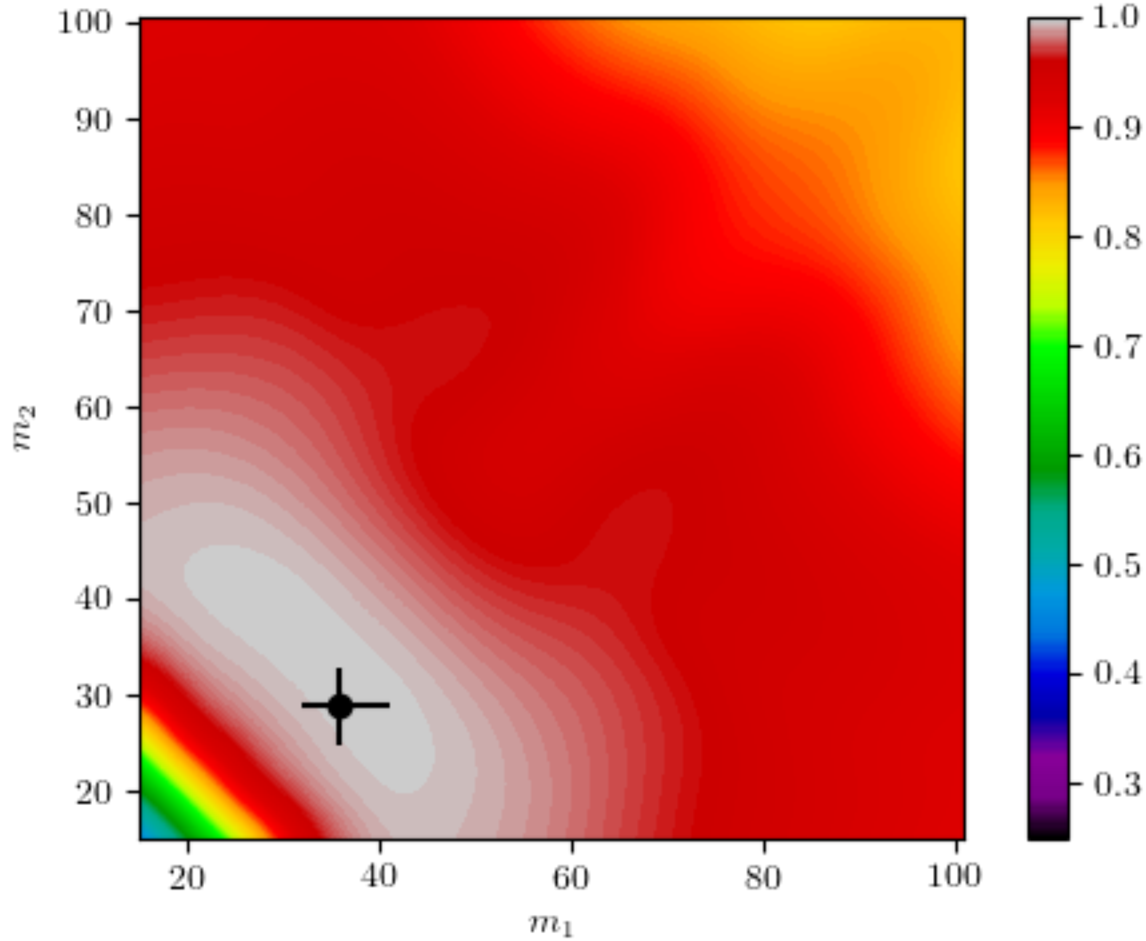
```
from pycbc.waveform import get_td_waveform

for apx in ['SEOBNRv2']:
    hp, hc = get_td_waveform(approximant=apx,
                             mass1=47.93,
                             mass2=36.59,
                             spin1z=0.9616,
                             spin2z=-0.89,
                             delta_t=1.0/16384,
                             f_lower=25)
```

There is an enormous degeneracy in the templates that makes it very hard (impossible) to determine masses with accuracy!

The use of the "secret template" does *not* change our results! LIGO knows this but will not admit it.



The Niels Bohr International Academy



LIGO's masses are *completely* uncertain!

This use of precalculated templates is a problem, Cornish concedes. "With a template search, you can only ever find what you're looking for." What's more, there are some templates, such as those representing the waves created by certain types of supernovae explosions, that LIGO researchers can't create.

And there are legitimate questions about that trust. *New Scientist* has learned, for instance, that the collaboration decided to publish data plots that were not derived from actual analysis. The paper on the first detection in *Physical Review Letters* used a data plot that was more "illustrative" than precise, says Cornish. Some of the results presented in that paper were not found using analysis algorithms, but were done "by eye".

Brown, part of the LIGO collaboration at the time, explains this as an attempt to provide a visual aid. "It was hand-tuned for pedagogical purposes." He says he regrets that the figure wasn't labelled to point this out.

The Niels Bohr International Academy

morphology

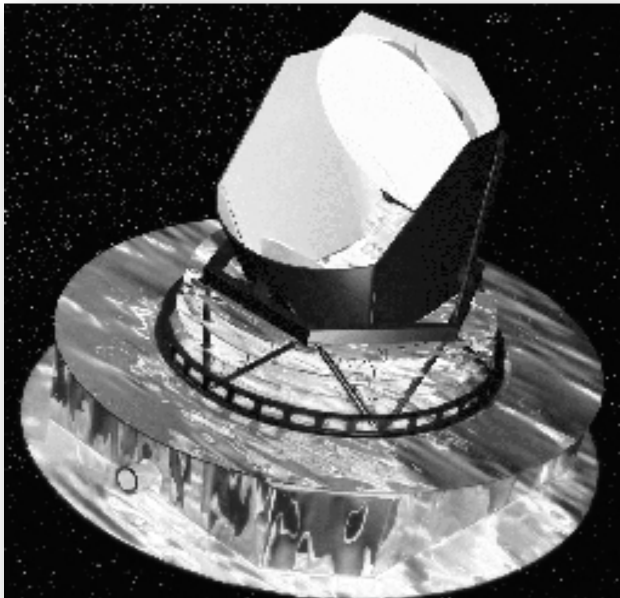
/mɔːˈfɒlədʒi/

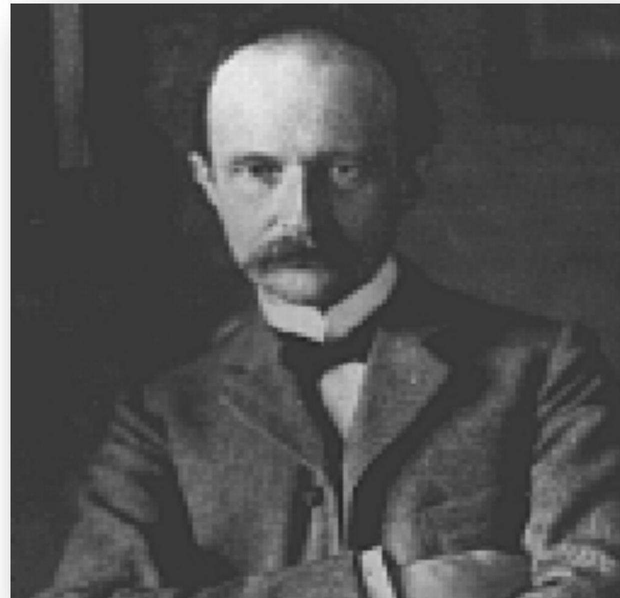
noun

1. the study of the forms of things, in particular:
2. a particular form, shape, or structure.

The Niels Bohr
International Academy

Planck × 2





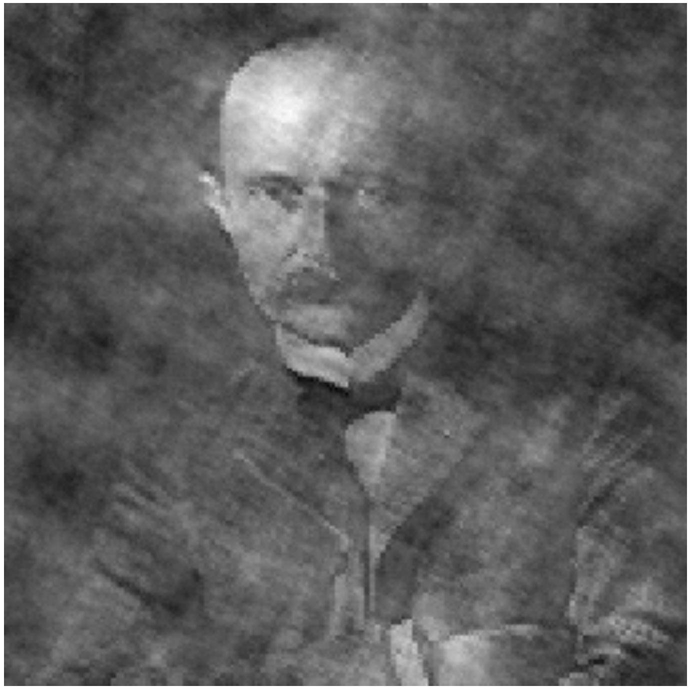
Any image can be written as

$$\sum A_k e^{i\varphi_k} \exp\left[i\mathbf{k}\cdot\mathbf{x}\right]$$

But what do we get if we mix the amplitude from the Planck satellite image with the phases of the Max Planck image?

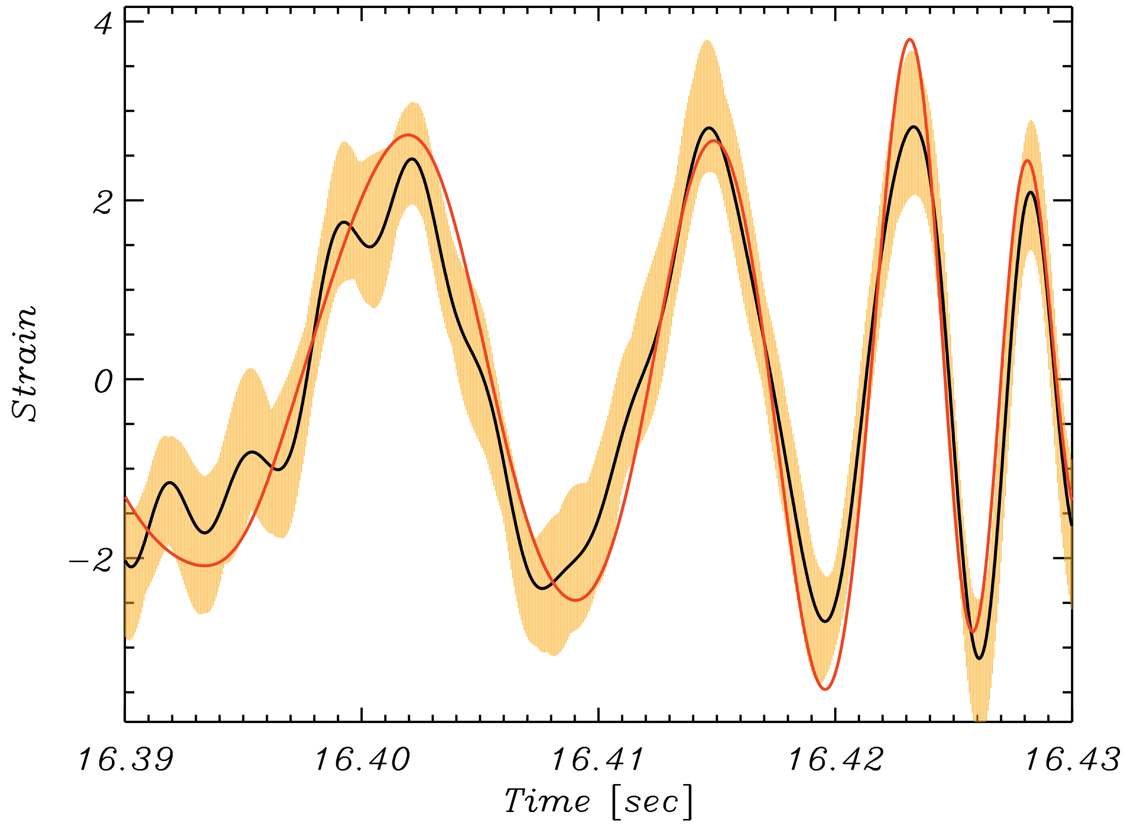
The Niels Bohr International Academy

Surprisingly, we get Planck himself!



Morphology is largely determined by phases and not amplitudes.
The power spectrum fails to tell the whole story.

Lung-Yih Chiang
http://www.asiaa.sinica.edu.tw/~lychiang/index/node8.html



The unbiased “best common signal”

The Niels Bohr International Academy

No.	Name	Old value	New value
1	ra (α)	1.5730257	2.1140412
2	dec (δ)	-1.2734810	-1.2518563
3	distance (d_L)	476.7564547	527.0627598
4	inclination (ι)	2.9132713	3.0021228
5	mass1 (m_1)	39.0257656	38.4900335
6	mass2 (m_2)	32.0625631	32.3104771
7	polarization (ψ)	5.9925231	2.6053099
8	spin1_a (a_1)	0.9767961	0.9635978
9	spin1_azimuthal (θ_1^a)	3.6036952	4.7164805
10	spin1_polar (θ_1^p)	1.6283548	1.9250337
11	spin2_a (a_2)	0.1887608	0.2894704
12	spin2_azimuthal (θ_2^a)	3.4359460	2.0230135
13	spin2_polar (θ_2^p)	2.4915268	0.7928019
14	tc (from 1126259462)	0.4175646	0.4151170
15a	coa_phase	0.6883212	N/A
15b	phase_shift ϕ_0	-0.9155276	1.7576289
15c	summed phase	-0.2272064	1.7576289
	SNR	24.36169	24.33653

SNR±1

Maximum likelihood methods are not useful! They cannot determine either the intrinsic features of a binary black hole system or determine its sky location!

The Niels Bohr International Academy

OBSERVATIONS:

- The use of templates can give a misleading estimate of the confidence of GW detection.
- Phase information should not be ignored.
- Ideally, GW detection should be based on the power of redundancy — including data from other sources.
- The fact that "GW" and "non-GW" components of GW150914 show a 7 ms time delay between Hanford and Livingston detectors has the unfortunate consequence that honest scientists can disagree about the interpretation of this event as a gravitational wave.

By the way, physicists should demand to see uncertainties and should know how they were calculated!

PHYSICAL REVIEW
OCTOBER 1, 1956
VOLUME 104, NUMBER 1

Question of Parity Conservation in Weak Interactions*

T. D. LEE, *Columbia University, New York, New York*

AND

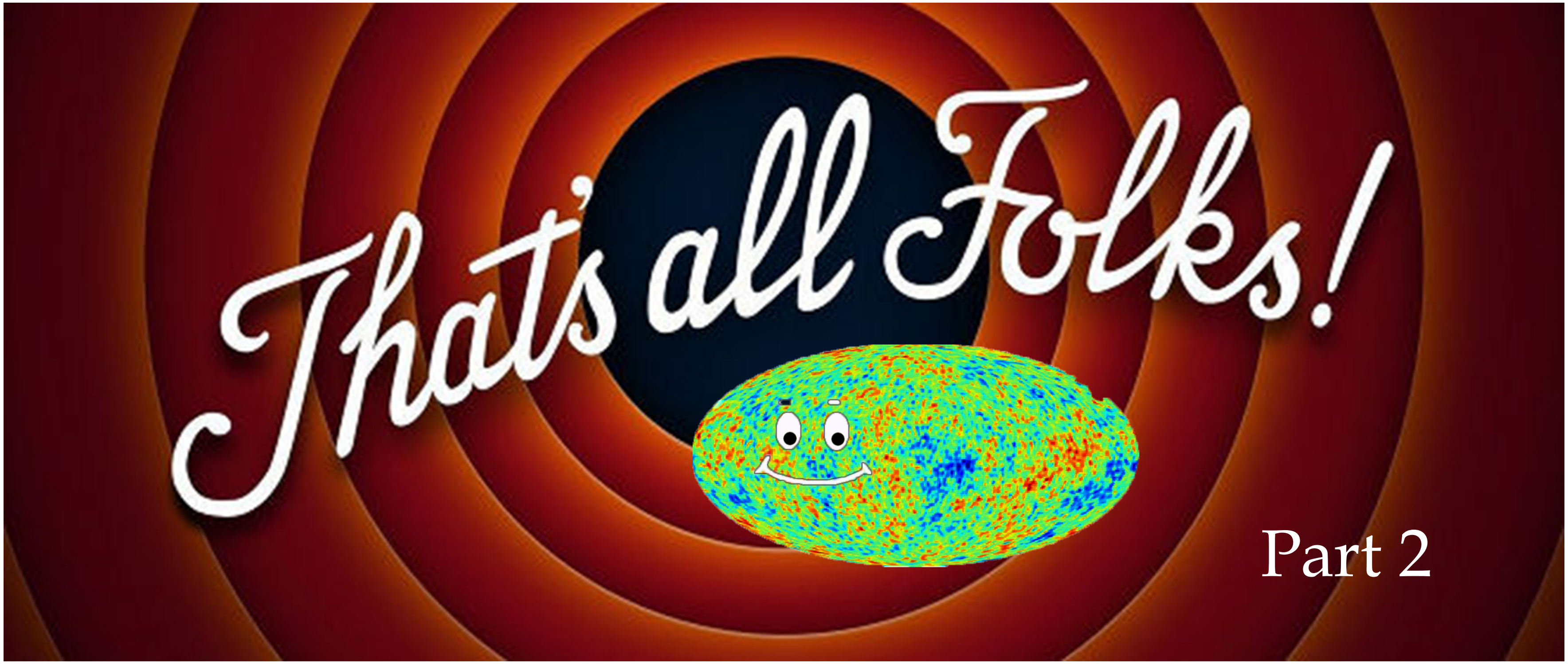
C. N. YANG,† *Brookhaven National Laboratory, Upton, New York*
(Received June 22, 1956)

The question of parity conservation in β decays and in hyperon and meson decays is examined. Possible experiments are suggested which might test parity conservation in these interactions.

RECENT experimental data indicate closely identical masses[1] and lifetimes[2] of the $\theta^+(\equiv K_{\pi 2}{}^+)$ and the $\tau^+(\equiv K_{\pi 3}{}^+)$ mesons. On the other hand, analyses[3] of the decay products of τ^+ strongly suggest on the grounds of angular momentum and parity conservation that the τ^+ and θ^+ are not the same particle. This poses a rather puzzling situation that has been extensively discussed.[4]

One way out of the difficulty is to assume that parity is not strictly conserved, so that θ^+ and τ^+ are two different decay modes of the same particle, which necessarily has a single mass value and a single lifetime. We wish to analyze this possibility in the present paper against the background of the existing experimental evidence of parity conservation. It will become clear that existing experiments do indicate parity conservation in strong and electromagnetic interactions to a high degree of accuracy, but that for the weak interactions (i.e., decay interactions for the mesons and hyperons, and various Fermi interactions) parity conservation is so far only an extrapolated hypothesis unsupported by experimental evidence.

The Niels Bohr International Academy



The Niels Bohr International Academy