LIGO- G040261-00-Z

Review of S2 BNS inspiral search

Alan Weinstein
(LIGO Laboratory / Caltech)

For the LSC Internal review committee:

Vicky Kalogera, Bill Kells, Alan Weinstein, John Whelan, Alan Wiseman

LIGO LSC Meeting

June 5, 2004

Tufts

LSC Inspiral Analysis Working Group
(Formerly known as the Inspiral Upper Limit Group)

Active members:

Bruce Allen [1], Stanislav Babak [2], Sukanta Bose [3], Patrick Brady [1*],
Duncan Brown [1], Alessandra Buonanno [4], Yanbei Chen [5],
Thomas Cokelaer [2], Nelson Christensen [6], Jolien Creighton [1],
Stephen Fairhurst [1], Gabriela González [7*], Gareth Jones [2],
Eirini Messaritaki [1], Brian O'Reilly [8], Ben Owen [9], Yi Pan [5],
Andy Rodriguez [7], B. Sathyaprakash [2], Peter Shawhan [5],
Michele Vallisneri [5], Darren Woods [1], Natalia Zotov [10]

Institutions:

* *Co-chairs*

[1] University of Wisconsin—Milwaukee, [2] Cardiff University,
[3] Washington State University, [4] Institut d'Astrophysique de Paris,
[5] California Institute of Technology, [6] Carleton College,
[7] Louisiana State University, [8] LIGO Livingston Observatory,
[9] Penn State University, [10] Louisiana Tech University

LSC Internal reviewers: Vicky Kalogera, Bill Kells, Alan Weinstein, John Whalen, Alan Wiseman

Inspiral searches in progress

- **BNS inspiral search: LIGO S2 *coincidence* analysis**
  - Neutron star binaries with component masses between 1 and 3 $M_{\odot}$
  - Coincident analysis: “L1 and (H1 or H2)”
  - Focus on detection rather than upper limit.
  - Still not in astrophysically interesting regime
  - Analysis and paper nearing completion
  - Reporting on this today
- **BNS inspiral search: Joint analysis of LIGO S2 + TAMA DT8**
  - analysis under intense development
- **BNS inspiral search: Search using LIGO+GEO S3 data**
  - analysis under intense development
- **Binary Black Hole MACHO Search using LIGO S2 & S3 data**
  - Binaries with component masses between 0.1 and 1 $M_{\odot}$
  - analysis rather advanced, plan to finalize by August LSC meeting
- **Search for Non-Spinning Binary Black Hole Systems**
  - analysis under intense development
- **Search for *Spinning* Binary Black Hole Systems**
  - analysis in relatively early stages of development

Review process and progress

- **The committee has focused, for the last ~4 months, on the S2 BNS search, mainly because the analysis is furthest along: a first draft of a paper was available before the last LSC meeting, and a rather mature (but not yet complete) draft is now under intense scrutiny by the committee.**
- **A long list of questions and concerns were formulated 4 months ago, and by now most or all of them have been addressed satisfactorily.**
- **A few important ones are still being worked on by the inspiral group, most notably the detailed analysis of systematic errors.**
- **We have had two marathon (3 hr) code reviews, and plan to have at least one more in the next week or two.**
- **The Inspiral notebooks have been very valuable tools, and the group has been pretty good about documenting their work.**
- **The committee is pleased with the progress. We expect to be able to give our blessings to a complete analysis and paper draft in the next few weeks.**

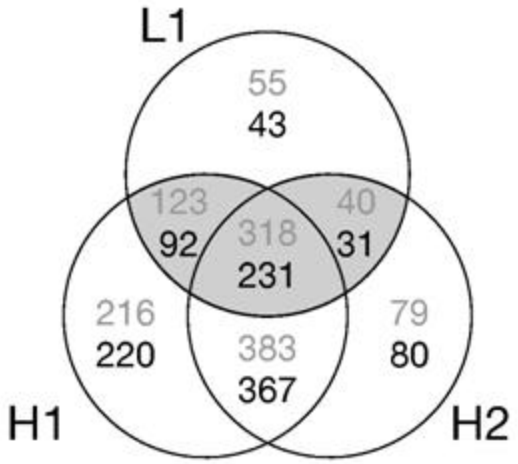
Differences between S1 & S2

- **The S2 BNS paper has several noteworthy features:**
  - extra-galactic reach of the 3 detectors
  - longer observation time
  - coincidence requirement for detection confidence (L1 and (H1 or H2))
  - estimation of background (accidental coincidence) from data
  - pipeline is largely automated to minimize bias

- **But the upper limit is only a bit better:**

Rate < ~50 per year per MWEG

c.f. **S1 result: < 170 per year per MWEG**



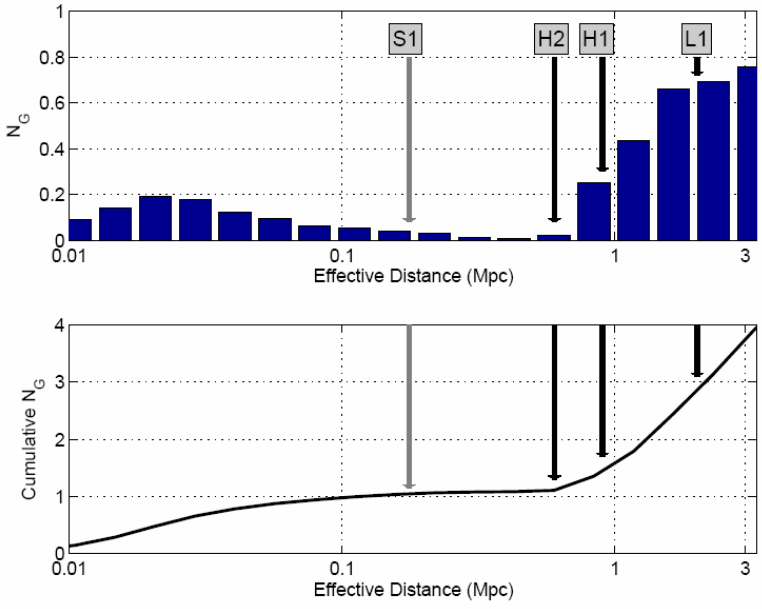
S2 Reach for Inspiral Searches

S1 visible range for 1.4+1.4 $M_{\odot}$
(optimally oriented, with SNR=8) **:**

L1 ~175 kpc ← Milky Way and Magellanic Clouds
H1 ~38 kpc ← Most of Milky Way
H2 ~35 kpc

S2 visible range for 1.4+1.4 $M_{\odot}$:

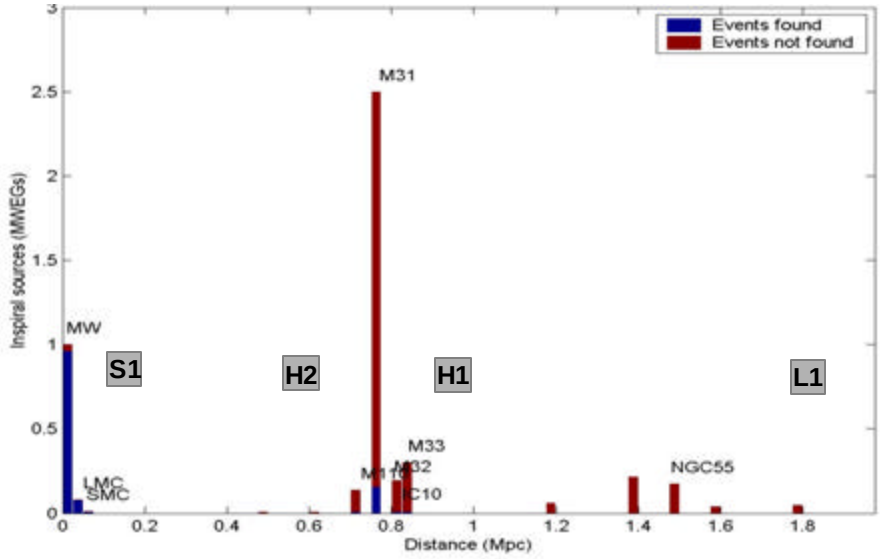
L1 ~1.8 Mpc ← Reaches M31, M32, M33, M110
H1 ~0.9 Mpc ← Barely reaches M31, etc.
H2 ~0.6 Mpc



• Significantly larger reach from S1 ➔ S2

• essentially 100% sensitive to sources in MW, LMC, SMC, and now sensitive to Andromeda (M31/M33);

• BUT, there's a lot of empty space between MW and M31!



Coincidence requirement: pros/cons

Require coincidence between LLO and LHO (H1 or H2 or both):

- **suppress noise trigger rate**
- **permit lower threshold**
- **estimate remaining background from data**
- **greatly increase detection confidence**

Down-sides:

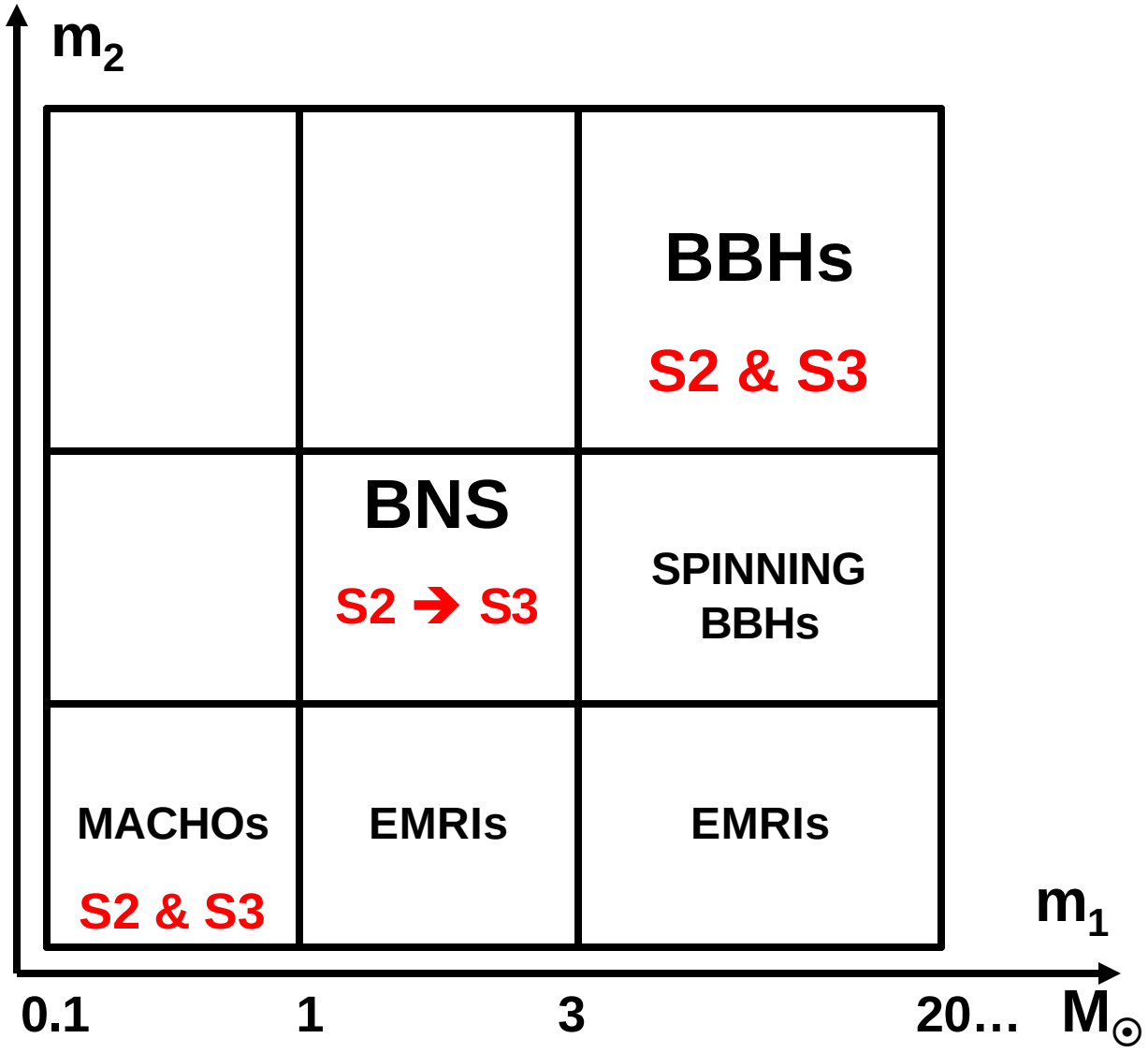
- **times when only a single detector/site is in science mode are excluded, reducing live time to 1/3 of total**
- **only as sensitive as the less-sensitive detector/site**

Optimized for detection confidence, NOT optimized for best upper limit

- **Use the single-IFO data, in coincidences with TAMA (and GEO for S3)**
  ➔ analysis in progress

- **The group has decided to analyze all science data for S3, and pursue detection with coincidences, upper limit with full data set.**

Binary mass space coverage



Binary mass space coverage

S2 BNS paper is furthest along,

but the *same* pipeline (different/more templates) is being used for the binary MACHO search (Duncan's thesis), and for the binary BH searches (non-spinning, spinning, EMR); so reviewing these works should be a straightforward extension of what the reviewers have covered so far.

Of course, new complexities appear with these other searches; we can concentrate on them, not the pipeline; we are being convinced that the pipeline is efficient and robust.

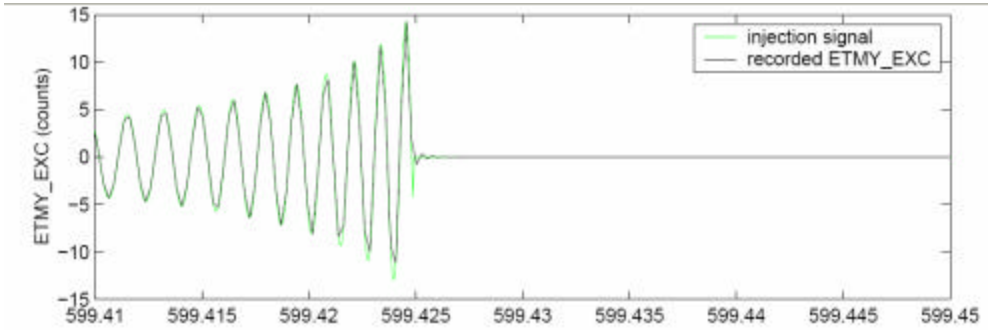
The aim of the current pipeline infrastructure is to be able to handle the general case of $m_1, m_2 = [0.1 \ldots 20] M_\odot$

The pipeline infrastructure can also accommodate any other short-duration matched filtering (merger waveforms, ringdowns…)

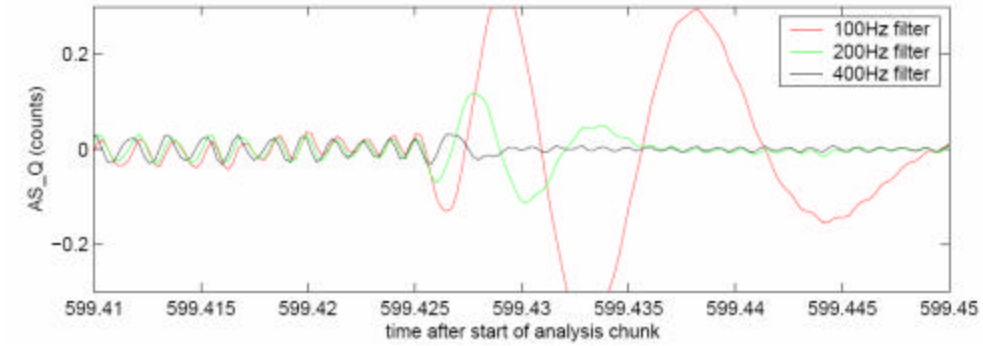
Issues addressed (1)

- **Astrophysical motivations: Are they adequately articulated in the paper? Non-GW sources of orbital decay truly negligible?**
  - Words (and references) will be added to current draft.
- **Models: Are the templates modeled correctly? Is the bank tiled sensibly?**
  - Many cross-checks, using different codes and approximations, have been performed; the template modeling seems to be robust.
- **Is the effect of spin and higher order terms truly negligible?**
  - Yes, for 1-3 $M_{\odot}$ BNSs; documents referenced in paper.
- **How does the presence of a merger waveform after the chirp affect the efficiency?**
  - Fortuitously, hardware chirp injections, which end abruptly, kick up the ETMs, causing a large broadband merger-like signal. These had no significant effect on the detection, merger time, SNR, effective distance…

Hardware chirps produce mergers!

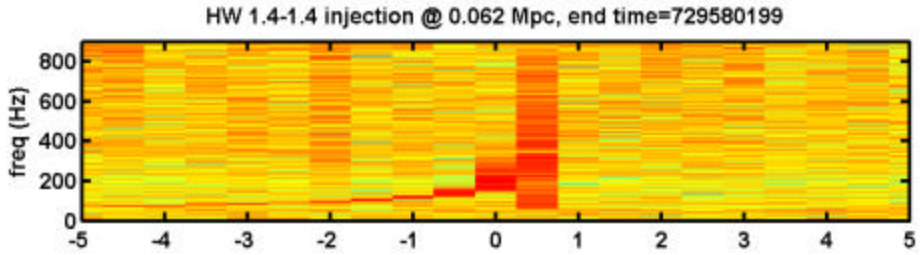


Injected waveform



Filtered AS_Q

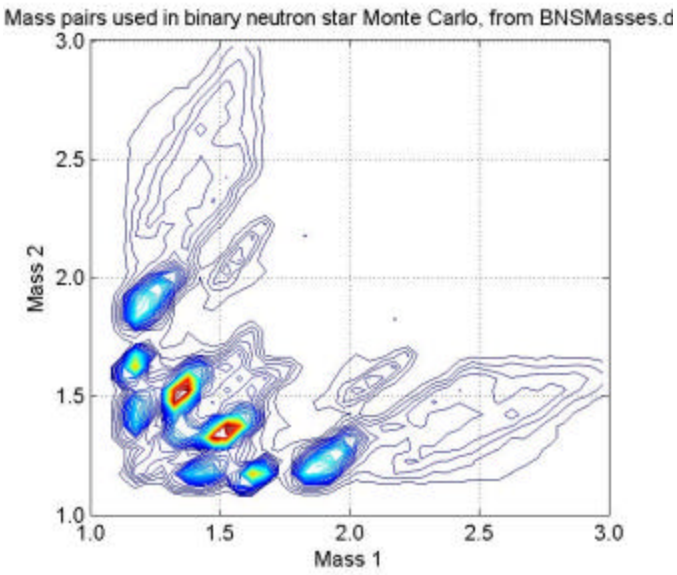
large ringing at 100-200 Hz, but also at 200-400 Hz

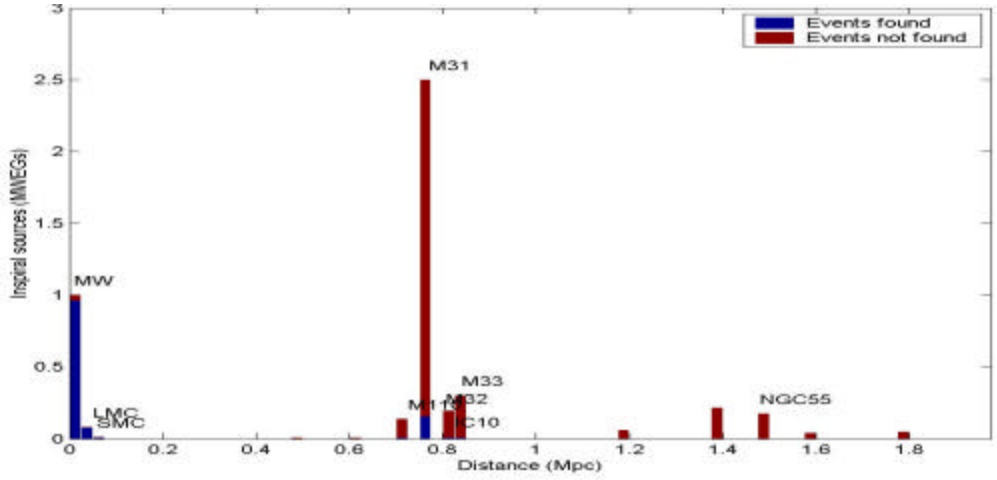


Spectrogram showing broadband burst

Issues addressed (2)

- **How is efficiency calculated? How does it depend on the model of the source spatial distribution, m_1/m_2 distribution?**
  - The uncertainties and assumptions made in the catalogs of nearby galaxies, and population synthesis model, are under close scrutiny, and these uncertainties are being propagated into the error on
    $N_G = \varepsilon N_{MWEG}$
    including correlations between distances and masses (BLL) of galaxies.
  - Plot efficiency vs effDistance, for smoother distribution of sources?





Coincidence Requirements

An "event candidate" is required to be detected by same template in L1 and in either H1 or H2

Consistency criteria depend on the detector pair

	H1-H2	L1-H1 / L1-H2
Time:	$\Delta t < 1$ ms	$\Delta t < 11$ ms
Template:	$\Delta m_1, \Delta m_2 = 0$	$\Delta m_1, \Delta m_2 = 0$
Effective distance:	$\frac{\lvert D_{H1} - D_{H2} \rvert}{D_{H1}} < 0.5 + \frac{2}{r_{H1}}$	No requirement, since LHO and LLO are not exactly co-aligned

MC simulation is best check that these cuts are efficient.

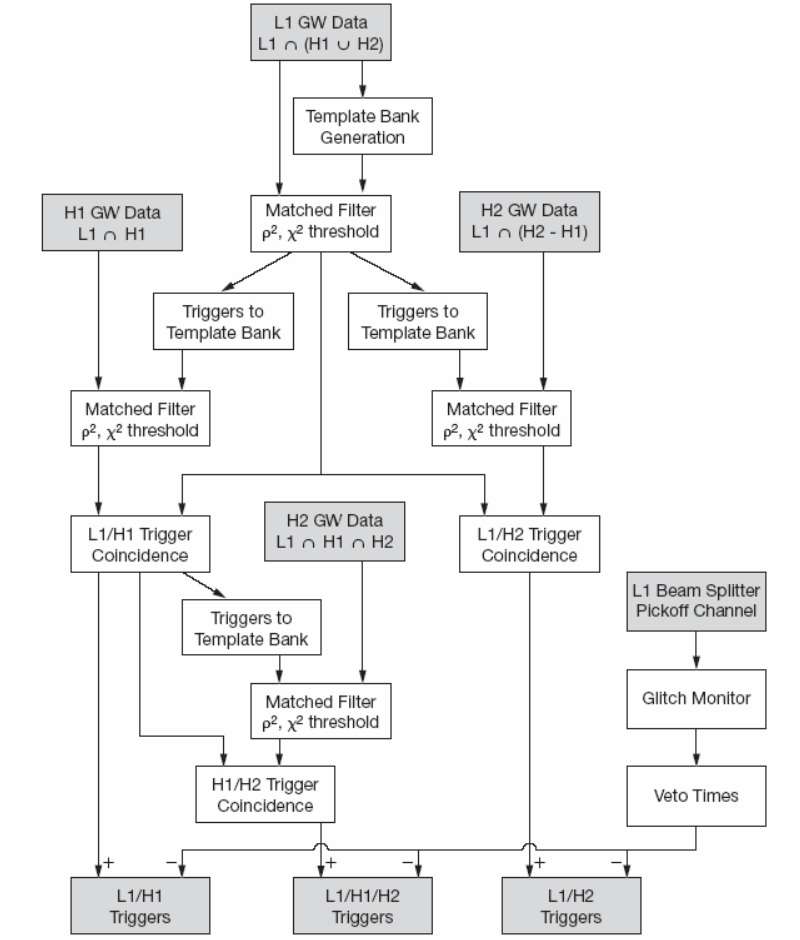
Also need plots of resolution on Δt, effDist, template Δm_1, Δm_2.
Non-zero time-slide search (as was done for burst review)?

Issues addressed (3)

- **Observation time: is it well understood? Effect of "chunks", "segments", chunk edges, overlaps, etc.**
  Data handling: what assurances / tests do we have that the observation time is correct?
  - MANY independent checks have been done, bugs found and fixed. The observation time for the search and for the simulations are understood to the second. However, this will all have to be re-done with the recent re-runthrough of the data pipeline, and the application of the L1:LSC-POB_I veto.
- **Granularity of PSD, calibration calculations. To what extent does the result depend on the averaging of the noise PSD and the calibration over "chunks"?**
  - Many checks have been done (eg, comparing hardware & software injections), at varying levels of rigor. The primary effect is on the effective distance, and this has a small effect on the overall efficiency (dominated by MW+LMC+SMC).
  - Using the median of the PSD rather than the mean makes it less sensitive to the presence of a signal or a glitch in the data stream.
- **What assurances do we have that the hardware and software injections have been passed through the same FULL pipeline as the search data?**
  - The group will treat us to another detailed code review to convince us that this is indeed the case, to a large extent.

Issues addressed (4)

- **Why is the pipeline "asymmetric", treating L1 differently from the other 2 detectors? Why can't it be "flat"?**
  - We have been convinced that for a coincidence analysis, it is completely unnecessary to filter H1 / H2 data with the full template bank; only the templates which trigger on the L1 data are needed.
    This saves HEAPS of CPU time, with no loss of generality.
  - The paper will contain language to make this clear, despite the complexity of the flowchart.
  - Still, this approach does not generalize easily (eg, to S3, where the noise curves are sufficiently different that the template bank tiling needs to be different).
  - For S3, all good data will be filtered with full template banks built for each detector.



Issues addressed (5)

- **Playground data used for tuning cuts: is it truly representative of the full data set?**
  - The playground duty cycle and SenseMon ranges and range distributions are statistically very similar to the full S2 dataset.
- **Is it necessary to use a playground sample at all? Can't delayed coincidences be used?**
  - Much discussion, differences in opinion.
- **Can detections be made in the playground data set?**
  - The playground data will not be used for the upper limit calculation, but *will* be used for identifying potential candidates for detection.
- **The upper limit on the event rate is calculated using the formula (2.303+lnP). Documentation? Reference?**
  - Peter Shawhan has verified that this formula produces sensible frequentist confidence levels.
  - Derivation to be published in proceedings of GWDAW (2003).
  - For the S2 paper, they will take the conservative approach of setting P (probability that all background lies below the loudest event) = 1. They will use the S2 paper as an opportunity to introduce the formula, but not use it.
  - They will evaluate P (it is around 0.8) but not its systematic error.
  - There is some hint that the time-delayed coincidences do not accurately represent the background; under study.

Background due to accidental coincidences

Estimate background due to accidental coincidences, using 37 time lags, each longer than the longest template (~ 4 seconds): ± (17 + 10n), n = 0…18.

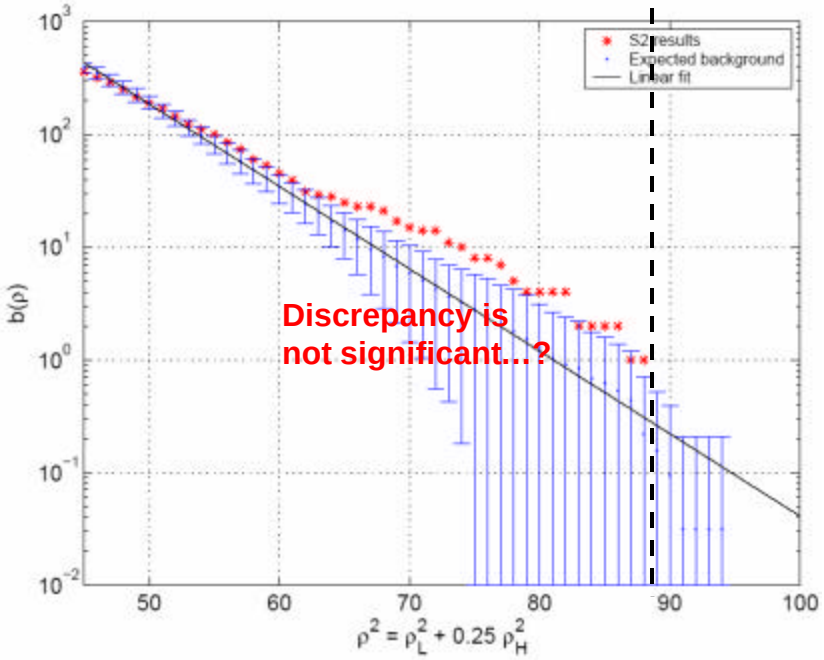
I : Number of accidental triggers vs r^2

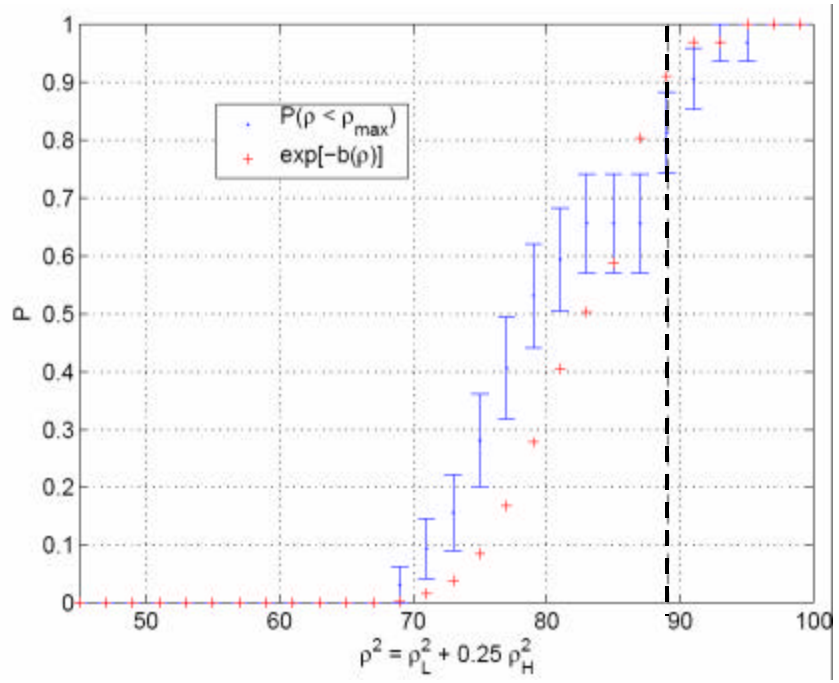
*** : Number of in-time triggers**

Probability that there are no background events above some maximum r^2

Loudest in-time coincident trigger in S2 has r^2 = 89;

20% chance that it is a background event.





Sources of Systematic Uncertainty

On e :

- Distances to galaxies
- Accuracy of template waveforms
- distribution of m_1 , m_2
- Calibration
- Effect of cuts on real vs. simulated signals
- Finite statistics of simulation

On N_{MWEG} :

- Number of sources in galaxies other than the Milky Way
  - Use blue light luminosity
  - Metallicity corrections
- Blue light luminosity of the Milky Way

Since the efficiency is evaluated with SAME pipeline as data,

with injections throughout S2 data sample,

there's no syst error associated with pipeline;

bugs or non-optimal pipeline algorithms can only lead to worse UL.

Simple sanity checks give ~ expected efficiency

Issues addressed (6)

- **The main difference between this paper and S1 is the use of coincidence for detection confidence. Is this analysis capable of making a detection? What procedures would be followed?**
  - Pipeline is automated, but …
  - The behavior of the detector is not ideal (non-Gaussian, non-stationary) and not yet fully understood
  - Establishing data quality cuts and aux-channel vetoes NOT yet automated
  - Detection confidence test gauntlet also NOT yet automated, or even fully designed; it may be impossible to design an unbiased set of tests and criteria before "opening the box"
  - Does that mean that we can't claim a detection if we have candidate(s)? No, only that detection confidence WILL be subject to bias...
  - As we build confidence in our detector characterization, we will build confidence in our ability to claim detection with minimal bias, and automate the confidence gauntlet as much as possible.

The group is carefully considering what should be said in the paper. Coincidence is only one (and not necessarily necessary) requirement for detection confidence.

Detection confidence

(Some) detection confidence tests for loudest triggers:

- Identify candidates with low background probability
- Check GW channel time series near candidate
- Is there a coincident candidate in burst search? Ringdown?
- What data quality flags were on>
- Was there anything strange in sci-mon or ops e-logs?
- Any data corruption evident between the data used in the analysis and the raw data archived at Caltech?
- Are injection channels clear?
- Are the candidates stable against changes in segmentation?
- Are the candidates stable against small changes in calibration consistent with systematic uncertainties?
- What are the parameters of the event? Are the masses reasonable? Is the distance reasonable? What position information is available via the time-delays? Are distances as measured in both instruments consistent with position information? Can the harmonics give useful information?
- If we cut signals by regions of parameter space, does this change the false alarm probability per week?
- What does the reconstructed waveform look like?
- Where does the candidate lie in parameter space of snr, chisq, masses, etc.
- Did the template bank ring-off all over? Is this consistent with a signal?
- How does the snr v time and chisq v time plot look?
- Make a follow up with coherent multi-detector code. How does it look?
- Are there are any auxiliary or PEM channels which indicate that the instrument was not behaving correctly?
- Lightning storms, high wind, other noisy weather?
- Are there any EM triggers in coincidence?
- Were any other detectors operational during the period when the candidate was identified?

Should this paper be published? Where should this paper be published?

- **This paper presents a new upper limit for BNS inspirals, 3× better than S1**
- **Still orders of magnitude away from astrophysical expectations**
- **Limit: 50 /yr/MWEG. Population synthesis: ~ 2 ×10^{-4} /yr/MWEG**
- **Nonetheless, it "expands our knowledge of (the absence of) GWs in the universe"**
- **Improved sensitivity: We've touched Andromeda!**
- **Presents a new method; building up foundations for future, deeper papers.**
- **Why not combine with S3 data?**
  - S3 is more sensitive, and will use the full data sample for upper limit – will blow S2 result away
  - S3 paper is still some months away; want to get this result out now, to lay foundations for S3 and beyond.
- **Why not combine with MACHO, BBH, etc, searches?**
  - Better to introduce each new search range or methodology with a separate paper, then combine methods for results from future, longer / more sensitive science runs
- **Will we be putting out a new paper for each science run?**
  - Well, why not? This is the bread and butter of LIGO's progress towards detection
  - Future science runs will be longer, we won't be submitting paper after paper in rapid succession
  - Each successive run should demonstrate that we are looking for a wider range of sources, looking deeper into the universe
  - We should be moving towards fewer, deeper papers, not a proliferation
- **Committee members share the desires of the inspiral group to publish this S2 BNS search as a stand-alone paper in the near future, then move on to S3 and broader mass ranges.**
- **An appropriate venue might be CQG.**

NEEDS

- **Much of the work of the Inspiral group is CPU-limited, even making maximal use of Beowulf clusters at UWM, CIT, PSU…**
  - Not so bad for BNS, but CPU resources limit the MACHO, BBH, and especially spinning BBH searches
- **Still, it is fair to say that the work and progress of the Inspiral group is fundamentally limited by manpower. There are many opportunities for grad students and postdocs to make big impacts.**
- **Otherwise, reviewers might be forced to do some work.**