Non-Parametric Distributional Tests for Gravitational Wave Transient Event Detection

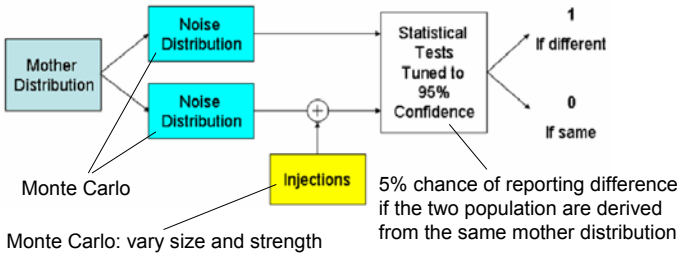
Yeming Shi[1], Erik Katsavounidis[1], Michele Zanolin[2]

[1] Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [2] Embry-Riddle Aeronautical University, Prescott, AZ 86301, USA

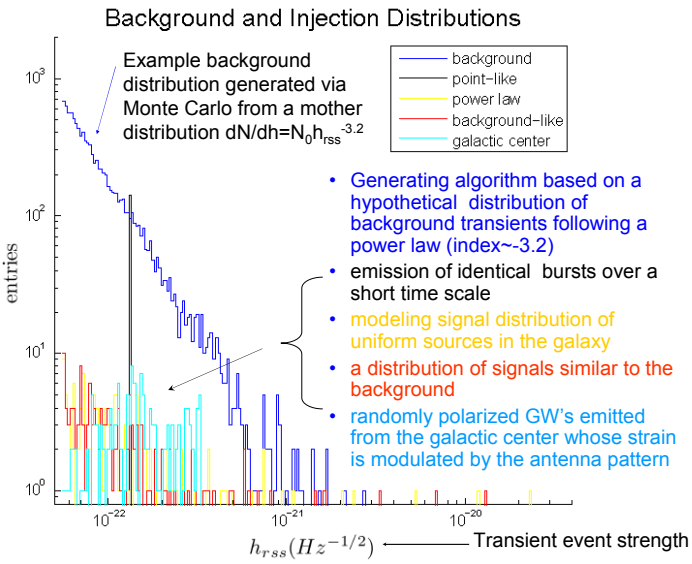
1: Motivation

- A counting-experiment approach has been traditionally used in searches for gravitational wave events where the significance of the observation is based on a comparison of the *number of events* between foreground and background
- A complimentary approach in searches for gravitational wave events can rely on a comparison of the *shape of distribution of event properties* among the two populations
- These distributional tests are very sensitive to the shape, though not so much to the size of the source population

3: Monte Carlo Experiment



4: Background and Injection Distributions



- Generating algorithm based on a hypothetical distribution of background transients following a power law (index~-3.2)
- emission of identical bursts over a short time scale
- modeling signal distribution of uniform sources in the galaxy
- a distribution of signals similar to the background
- randomly polarized GW's emitted from the galactic center whose strain is modulated by the antenna pattern

6: Comparison to Counting at Low FAR

- In a counting experiment with a <<1 expected background events, say, 0.1, according to the Feldman & Cousins formulism one would need 2 events in the foreground to claim detection at the 95% CL.
- **A distributional test, which requires higher statistics in order to perform a sensible comparison of distributions, beats the counting method at low FAR if it reaches a sizable efficiency when the criteria above is not met. This always happen for the cases considered in this analysis.**

7: Future Work

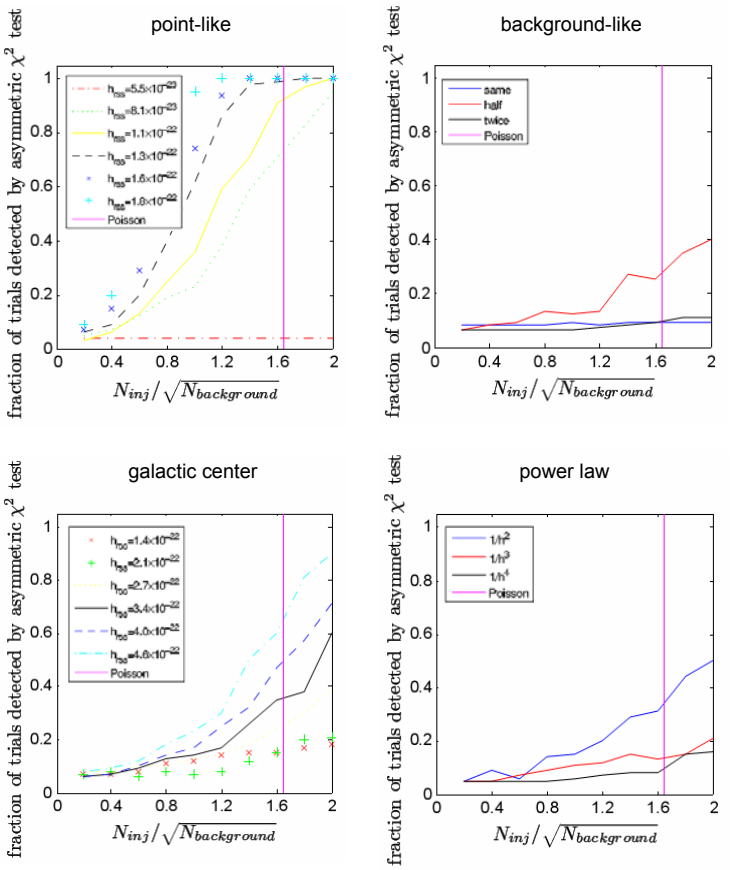
- Tune the method with background/simulated events involving real data
- Applications of the tests to real data as part of a search: with a given C.L., the tests can be used to set an upper bound on the flux of gravitational waves of a certain waveform.

2: Distributional Tests vs. Counting Statistics

- Several distributional tests were implemented and tested with simulated data. These tests do not assume analytical distributions of the background. In each test two distributions are compared, one corresponding to background and one to background+signal. A scalar measure of the discrepancy between the two distributions is computed, and if its value is outside the confidence interval, the test rejects the null hypothesis.
- Kolmogorov Smirnov : maximal difference in cumulative fraction functions
- Mann Whitney : for each population, the total rankings of its entries in the combined population
- χ^2 : the statistic is based on data binning and counting in each bin
- Asymmetric χ^2 : another implementation of the χ^2, assuming zero fluctuation in the background
- Counting statistics: we use the Feldman & Cousins approach which allows a counting experiment to exclude zero signal flux. Depending on threshold used, counting experiments can have variable false alarm rates. Here we study two cases of almost zero and large FAR.

5: Efficiency of the Tests

- Efficiency: fraction of injection scenarios with successful detection.
- The asymmetric χ^2 test is found to be the most sensitive. (χ^2 test the second most)
- Different tests are consistent with each other on the trials for which they detect the presence of injections.
- Plots: comparison between the tests and counting at large FAR.



- **For large FAR, distributional tests can perform better than counting method. However, the size of the injection necessary for detection at a certain efficiency depends on the strength and distribution of the GW's.**

Y. Shi (yeming@mit.edu)/M. Zanolin (zanolinm@erau.edu)/E. Katsavounidis (kats@ligomit.edu) — 12th GWDAW, MIT Cambridge MA USA LIGO G070884-00-0

The authors gratefully acknowledge the support of MIT's Undergraduate Research Opportunity Program (UROP) and of the LIGO Laboratory. We also acknowledge the input from and discussions within the LIGO Scientific Collaboration. This work was supported from MIT direct UROP funding.