Taming Instrument Glitches and Detecting Burst Signals

Paul T. Baker, Neil Cornish, Tyson Littenberg

Department of Physics, Montana State University

Abstract

With a network of gravitational wave detectors it is possible to distinguish between instrument artifacts (or glitches) and un-modeled gravitational wave signals. The LIGO-Virgo Burst group has developed several effective algorithms for detecting un-modeled, short duration burst signals such as might be generated by core collapse supernovae or associated with gamma ray bursts. Here we present a new algorithm that uses Bayesian model selection to decide if features in the data are better described as gravitational wave signals or instrument glitches. As a by-product, even when no burst signals are detected, this procedure produces cleaned data streams that are free of loud glitches. The cleaned data can also be used by standard template based searches for modeled signals such as binary inspirals, but now with significantly reduced backgrounds, making it possible to detect weaker signals.

Analysis Overview

Our analysis is conducted using a Markov Chain Monte Carlo (MCMC) search of simulated LIGO-Virgo data in the wavelet domain[1].

Our data

- 3 interferometers (IFO): Hanford 4km, Livingston, Virgo
- colored gaussian noise (S5 noise curves)
- "glitch": random amplitude spike
  (sine-gaussian or high sigma gaussian fluctuation)
- signal: coherent sine-gaussian GW burst

The data can contain any combination of noise, glitches, and signal. We analyze 16 seconds at a time and cover the frequency range from 32 to 512 Hz.

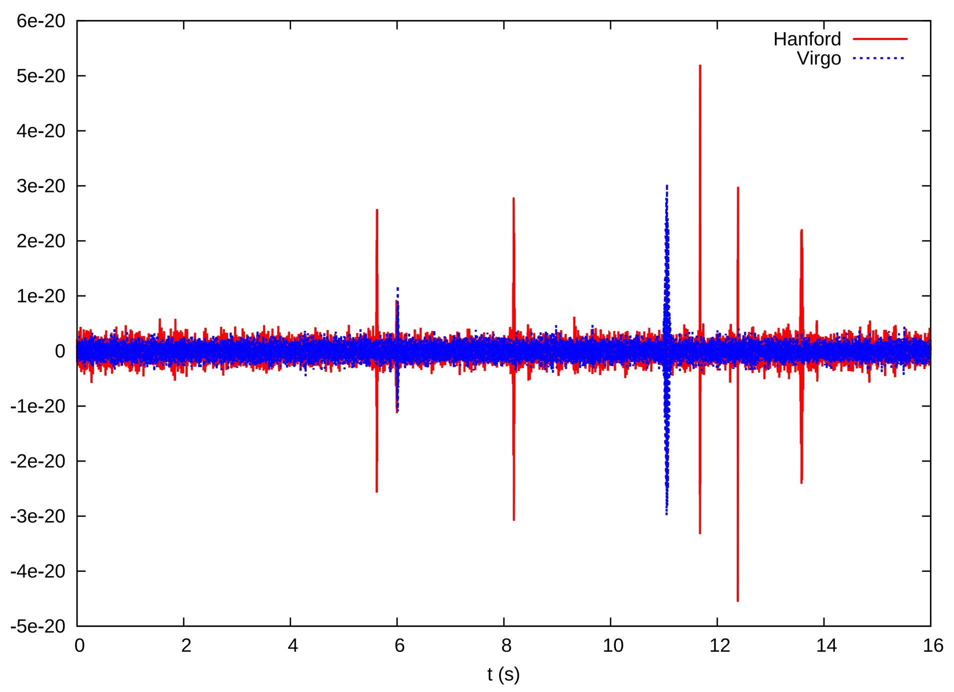

Figure 1: A glitchy simulated datastream (transformed back to the time domain) in 2 IFOs. The injected GW signal is at ∼6s. Other loud spikes are glitches.

The Glitch Model

We model the glitches as excess power in wavelet pixels. Each of the three IFOs has its own glitch model. The amplitude of each pixel on the wavelet grid is potentially a model parameter. The MCMC determines both how many pixels are non-zero and the value of those non-zero amplitudes.

Contact Information

Paul T. Baker
PhD Student

email: pbaker@physics.montana.edu
address: Montana State University
Department of Physics
Bozeman, Mt 59715

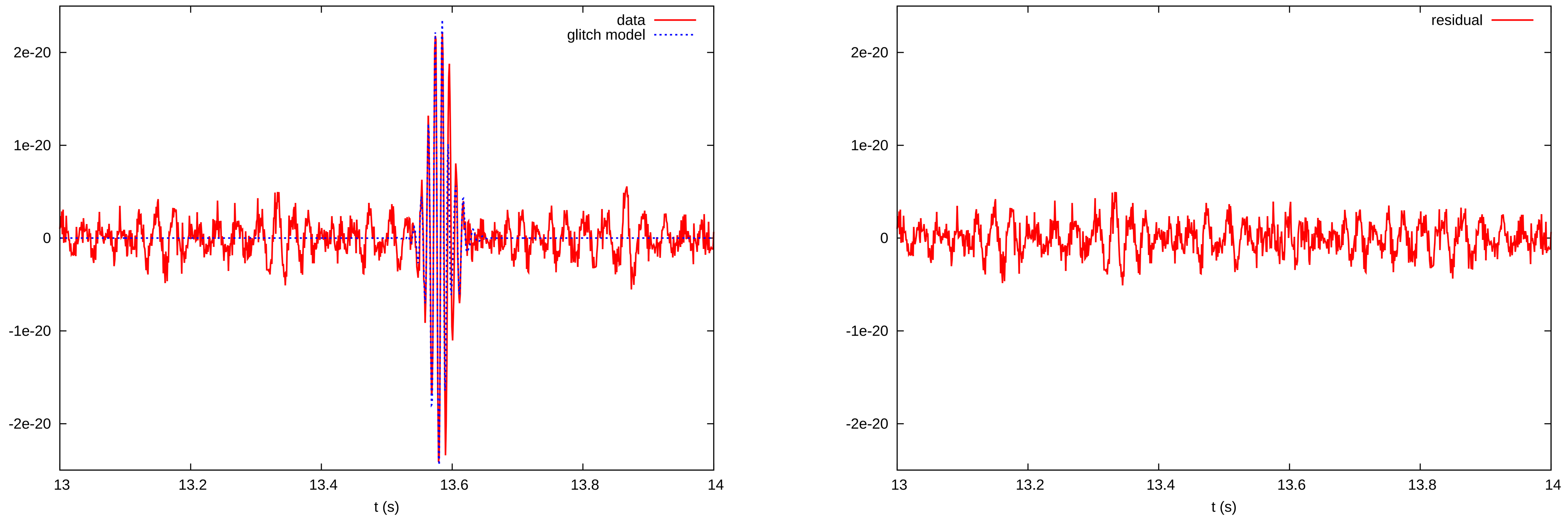

Figure 2: A time domain look at a simulated glitch in one IFO with the best fit glitch model superimposed. Also the same data after subtracting the best fit wavelets.

The Signal Model

- collection of wavelet pixels (potentially 16 384 params)
- one signal as seen at geocenter
- plus $(+)$ and cross $(\times)$ polarizations
  can target linear, circular or elliptical $(+1$ param)
  unpolarized needs two wavelet grids $(\times 2$ pixels)
- one sky location of origin $(+2$ params)
- geocenter signal is projected onto detectors

A technical note

The time shifting and polarization relations are easily conducted in the time and frequency domains respectively. To avoid costly wavelet transforms during run time, we construct look-up tables in advance, telling how individual wavelet pixels respond to such operations.

Glitch and Signal Together

The glitch and signal searches run simultaneously, both looking for excess power. The trick is that using fewer pixels leads to a worse fit to the data, but our prior belief in a model with few pixels is greater.

The posterior of the data, s, given model, $\mathcal{M}$, is the product of the likelihood of s given $\mathcal{M}$ and the prior belief in $\mathcal{M}$.

$$P(s, \mathcal{M}) = p(s|\mathcal{M})\, p(\mathcal{M}) \qquad (1)$$

The signal model can exploit this by using one wavelet pixel to represent excess power in all three IFOs, while the glitch model would need three.

Data Cleaning

For marginal detections we can "clean" the data for use by template based analyses that perform better on low SNR signals.

- run search to determine best fit coherent signal and glitches
- subtract best fit glitches from initial data stream
- rerun using different analysis technique

Types of Moves

Signal & glitch moves

amplitude move: propose new amplitudes for all pixels by a gaussian jump.

add pixel: propose to add a new pixel from a nearest neighbor histogram (see Figure 4)

remove pixel: propose to remove an existing pixel from a nearest neighbor histogram (set amp $= 0$).

Signal only moves

sky move: propose a new sky location of origin for the signal

polarization move: propose a new polarization angle or scale (linear or elliptically polarized signals only)

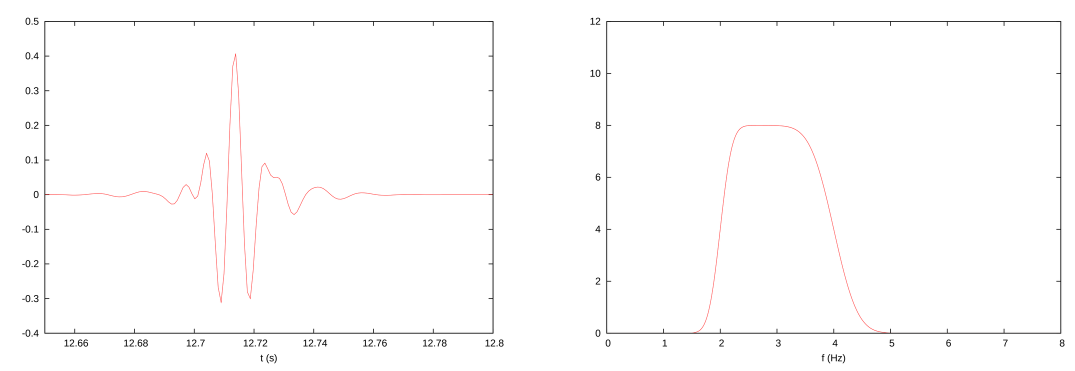

Figure 3: A wavelet pixel in the time and frequency domain. Wavelets are compact in both time and frequency, making them ideal for modeling both our signals and glitches

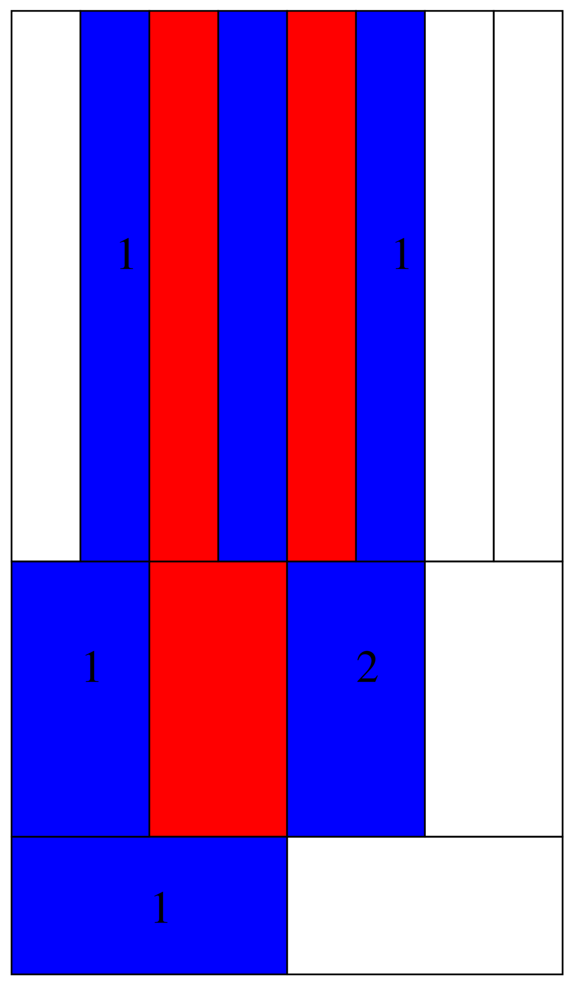

Figure 4: A didactic wavelet grid displaying the nearest neighbor histogram. Red pixels are non-zero, their neighbors are blue.

Targeted searches

- already look for clusters of pixels with nearest neighbor proposals
- add prior to favor clustered pixels
  e.g. $p(x) = \frac{n+1}{N+1}$ favors all neighbors equally
- choose shape of cluster to target specific waveforms
  BH ringdown: only left/right neighbors (const. frequency)
  BH inspiral: right-up and left-down (chirping signals)

The Bayesian Approach

Our search is the Bayesian analog of Coherent Waveburst (CWb)[2][3]. Where CWb uses regularization techniques to conduct unstable matrix inversions, we set constraints using prior knowledge.

Searle[4] showed that one method used by CWb to keep arbitrarily loud signals being found on the detetectors' "blind spots" is equivalent to a gaussian prior on signal amplitude:

$$p(x) \propto e^{-\frac{1}{2}\left(\frac{x}{\alpha\sigma}\right)^2} \qquad (2)$$

σ: standard deviation of the noise

$\alpha > 1$: a tunable constant, smaller $\alpha \implies$ greater penalty.

- signals proposed in blind spots have no effect on the likelihood
- prior will penalize signals for being too loud
- equivalent to saying: "we do not expect any really loud GWs from anywhere on the sky" (physically true)

The detection statistic

The use of a an MCMC allows for easy calculations of Bayesian evidence and therefore the Bayes factor. The Bayes factor tells us how much more likely it is that the data came from the first model relative to the second. It is essentially a betting odds. There is no need to devise and tune a detection statistic.

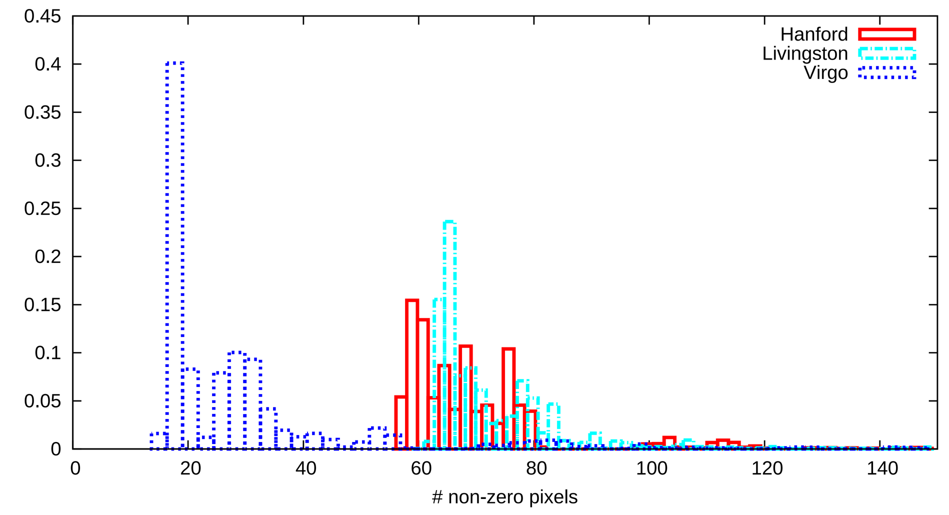

Figure 5: a Reverse Jump histogram showing the number of chain iterations in each non-zero pixel number model. All IFOs clearly show existence of glitches in the data

Future work

- finalize code, implement binary wavelet decomposition
- apples to apples comparison with other LIGO burst searches (CWb, Ω-pipeline)
- run on real data . . . **detect GWs** . . .

References

[1] T. Littenberg & N. Cornish, arXiv:1008.1577v1 [gr-qc]

[2] S. Klimenko, S. Mohanty, G. Mitselmakher, arXiv:0508068v2 [gr-qc]

[3] S.Klimenko, I.Yakushin, A.Mercer, G.Mitselmakher, arXiv:0802.3232v2 [gr-qc]

[4] A. C. Searle, P. J. Sutton, M. Tinto, arXiv:0809.2809v2 [gr-qc]

This work was supported by the National Science Foundation, and is part of ongoing work with the LIGO-Virgo Collaboration.