7-12-2014

Detection and estimation of unmodeled narrowband nonstationary signals: application of particle swarm optimization in gravitational wave data analysis

Soumya Mohanty

Detection and Estimation of Unmodeled Narrowband Nonstationary Signals

Application of Particle Swarm Optimization in Gravitational Wave data analysis

Soumya D. Mohanty
University of Texas at Brownsville
One West University Blvd.
Brownsville, Texas, 78520
mohanty@phys.utb.edu

ABSTRACT

The extraction of weak signals from instrumental noise is a critical task in ongoing searches for gravitational waves. A detection and estimation method, made feasible by Particle Swarm Optimization, is presented for a particularly challenging class of signals expected from astrophysical sources.

Categories and Subject Descriptors

G.1.1. [**Interpolation**]: Smoothing; G.1.2 [**Approximation**]: Least squares approximation; G.1.6 [**Optimization**]: Global Optimization; G.3 [**Probability and Statistics**]: Robust Regression; J.2 [**Physical Sciences and Engineering**]: Physics

Keywords

Gravitational Wave; Particle Swarm Optimization; Detection; Estimation; Regression; Spline; Non-stationary signal

1. UNNT SIGNALS, METHOD AND RESULTS

Let $\overline{x} = (x(t_0), x(t_0+\Delta), \ldots, x(t_0+(N-1)\Delta)) \in \mathbb{R}^N$, with Δ constant, represent a segment of gravitational wave (GW) detector output. (See [6] for a review of GW astronomy.) We have $\overline{x} = \overline{s} + \overline{n}$ and $\overline{x} = \overline{n}$, where $\overline{n}$ denotes a noise realization, in the presence and absence respectively of a GW signal $\overline{s}$. In the following, $\overline{n}$ is drawn from a zero-mean white Gaussian noise process with unit variance (WGN).

Fig. 1 shows the signals considered in this work. One is the s11WW signal taken from a catalog of core-collapse supernova simulations [7]. The other, which we call the 4-Peaks signal, is our *ad hoc* model for a signal from a long-lived rotational instability. The signals have been normalized to have a matched filtering signal to noise ratio (SNR) of 10, where $\mathrm{SNR}^2 = \|\overline{s}\|^2 = \overline{s}\overline{s}^T$. Due to the narrowband and possibly nonstationary features in their spectrograms and the fact that their waveforms are unpredictable for real sources,

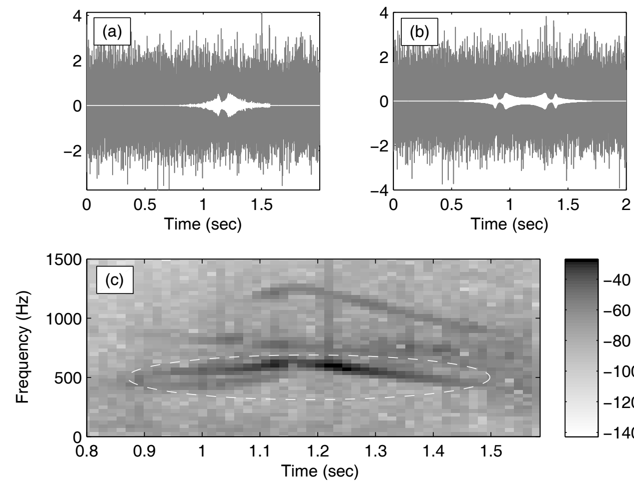

Figure 1: Panel (a) and (b) show (in white) the s11WW and 4-Peaks signal respectively superimposed on data realizations containing the signals (in gray). Panel (c) shows the spectrogram of the s11WW signal in dB. The dashed white ellipse shows a narrowband nonstationary feature.

such signals exemplify what we call Unmodeled Narrowband Nonstationary Transient (UNNT) signals.

Motivated by the slow variation of their amplitude envelope and instantaneous frequency relative to the instantaneous period of their carrier signal, we model UNNT signals as follows. (Here, $y_k \equiv y(t_0 + k\Delta)$ for any $\overline{y} \in \mathbb{R}^N$.)

$$s_k(\alpha, \gamma, \omega) \; = \; \sum_{j=0}^{M-3} \alpha_j \mathcal{B}_{j,k}(\gamma) \cos\left(\int_{t_0}^{t_k} dt' f(t'; \omega) \right) \;, (1)$$

where $\overline{\mathcal{B}}_j(t; \gamma)$, $j = 0, 1, \ldots, M-3$ are the B-spline basis functions [1] that span the space of cubic splines defined by a given set of knots $\gamma = (t_0, \gamma_1, \ldots, \gamma_M, t_{N-1})$, $M \ll N$. $f(t; \omega)$ is a linear spline with $K \ll N$ knots, with ω denoting the knots and the corresponding frequencies ($\in [0, f_{\max}]$).

We maximize the penalized log-likelihood ratio [2]

$$\mathrm{LLR}(\alpha, \gamma, \omega; \overline{x}) \; = \; \|\overline{x}\|^2 - \|\overline{x} - \overline{s}(\alpha, \gamma, \omega)\|^2 + \lambda \sum_{j=0}^{M-3} \alpha_j^2 \;, (2)$$

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.
ACM 978-1-4503-2881-4/14/07.
http://dx.doi.org/10.1145/2598394.2598439.

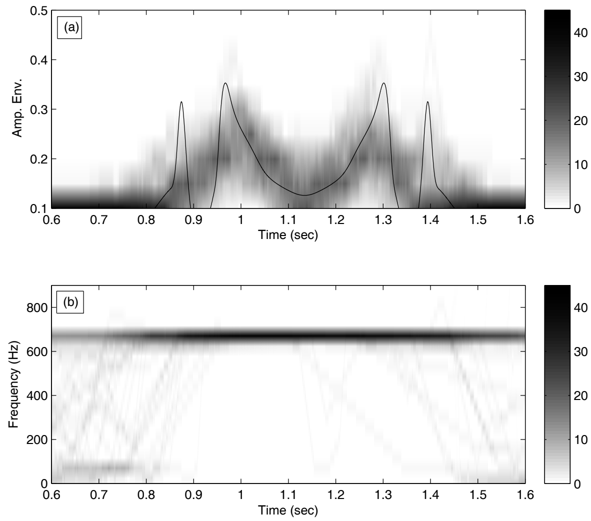

Figure 2: Histograms of (a) the estimated amplitude envelope, and (b) the instantaneous frequency at each instant of time for the 4-peaks signal. The true amplitude envelope is shown as the solid curve.

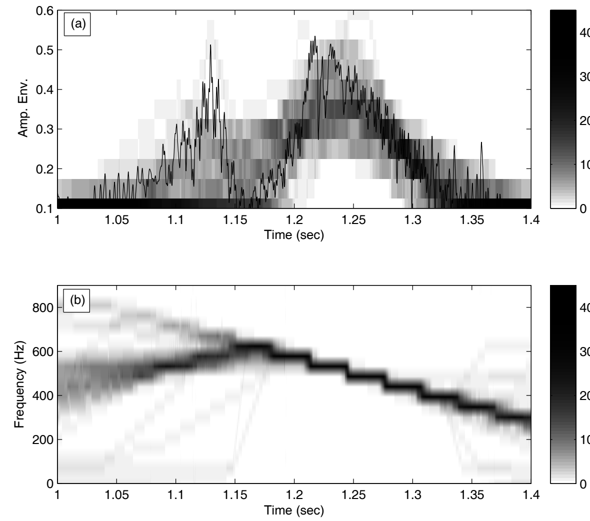

Figure 3: Histograms of (a) the estimated amplitude envelope, and (b) the instantaneous frequency at each instant of time for the s11WW signal. The true amplitude envelope is shown as the solid curve.

(i) analytically over α, under the constraints [3] $\alpha_j \geq 0$, $j = 0, \ldots, M-3$, and then (ii) numerically, using PSO [4], over γ and ω. Prior to step (ii), λ is determined using generalized cross validation [5]. Then $D(\overline{x}) = \text{LLR}(\widehat{\alpha}, \widehat{\gamma}, \widehat{\omega}; \overline{x}) = \max_{\alpha,\gamma,\omega} \text{LLR}(\alpha, \gamma, \omega; \overline{x})$ is the detection statistic and $\widehat{\alpha}$, $\widehat{\gamma}$ and $\widehat{\omega}$ are the point estimates of the signal parameters.

Simulation results are obtained for the following set up. *Data*: $N = 16384$; $\Delta = 1/8192$ sec. *UNNT model*: $f_{\max} = 800$ Hz, $M = 10$ and $K = 5$. *PSO*: 40 particles; 800 iterations; linearly decaying inertia from 0.9 to 0.4; ring topology with neighborhoods of 3 particles; Particle velocity along each coordinate $\leq 1/5$ of the coordinate range; both the acceleration constants = 2.0; 'let them fly' boundary condition. The dimensionality of the fitness function to be optimized by PSO is $M + 2K = 20$.

$D(\overline{x})$ is evaluated for the case $\overline{x} = \overline{n}$ with 100 independent realizations of $\overline{n}$, and for $\overline{x} = \overline{s} + \overline{n}$ with 45 independent realizations of $\overline{n}$ for each of the two signals (SNR=10 for both) . The effective SNR of the method, defined as

$$(\langle D(\overline{s} + \overline{n}) \rangle - \langle D(\overline{n}) \rangle)/\text{stdev}(D(\overline{n})) , \quad (3)$$

where $\langle X \rangle$ is the sample mean of a random variable X and $\text{stdev}(X)$ its sample standard deviation, is found to be 14.1 and 9.8 for the 4-Peaks and s11WW signals respectively. Since these values are comparable to the SNR for matched filtering, the method is seen to perform well in the detection of a wide range of signal morphologies without requiring any change in its settings.

Figs. 2 and 3 show the estimated amplitude envelopes and instantaneous frequencies for the case $\overline{x} = \overline{s} + \overline{n}$. In the time interval containing the signal, the instantaneous frequency estimates are tightly clustered around the true values. (The 4-Peaks signal has a constant instantaneous frequency of 650 Hz.) In line with the Fisher matrix for monochromatic signals, the estimation error is higher for the amplitude envelope than the instantaneous frequency. Nonetheless, information about the main peaks in the ampltiude envelope is retrievable despite the strong dominance of noise over the signal in the time domain (c.f., Fig. 1).

The results show that the method is promising. It is important to note that, due to the high-dimensional optimization involved, the method would be computationally prohibitive and, hence, infeasible without PSO.

2. ACKNOWLEDGEMENTS

This work is supported by NSF awards PHY-1205585 and HRD-0734800.

3. REFERENCES

[1] C. De Boor. *A practical guide to splines*, volume 27. Springer-Verlag New York, 1978.

[2] D. Ruppert, M. P. Wand, and R. J. Carroll. *Semiparametric regression*, volume 12. Cambridge University Press, 2003.

[3] D. Fraser and H. Massam. A mixed primal-dual bases algorithm for regression under inequality constraints. *Scandinavian Journal of Statistics*, pages 65–74, 1989.

[4] D. Bratton and J. Kennedy. Defining a standard for particle swarm optimization. In *Swarm Intelligence Symposium, 2007. SIS 2007. IEEE*, pages 120–127. IEEE, 2007.

[5] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.

[6] S. A. Hughes. Listening to the universe with gravitational-wave astronomy. *Annals of Physics*, 303(1):142 – 178, 2003.

[7] C. D. Ott. stellarcollapse.org, 2013. [Online; accessed 29-January-2014].