Preprint typeset using LaTeX style emulateapj v. 6/22/04

Possibilities for Processing Arecibo PALFA Data with Einstein@Home Computing Resources

Bruce Allen, AEI Hannover
Jim Cordes, Cornell

March 7, 2008
Version 1.00

Subject headings:

Summary

This document is based on notes of conversations between Jim Cordes and Bruce Allen at the Boston GWDAW meeting in December 2007 and in subsequent telephone conversations. The discussion concerns the prospects for analyzing pulsar data taken under the Pulsar-ALFA survey at Arecibo[1] by the Einstein@home community. The basic idea is to process a data unit consisting of a dedispersed time series to search for compact binary pulsars as a function of spin period, duty cycle, and orbital parameters. One approach is to parameterize orbits using acceleration and jerk values, corresponding to the first and second time derivatives of the apparent pulse frequency. Another approach is to use circular orbits and an economical grid of parameter values. Candidate signals will be returned to the central archive as a function of these four parameters and classified and mined.

PALFA Consortium

The Pulsar ALFA Project was proposed for and is managed by the PALFA Consortium, consisting of about 40 researchers (including students) at about 10 institutions around the world, concentrated in North America. Presently we are sampling the Galactic plane (±5 deg from the midplane) for the portions that can be sampled given the 20 deg zenith angle limitation on the telescope pointing. We plan to extend the survey to higher Galactic latitudes as a means for increasing the yield to millisecond pulsars, whose detection is especially inhibited by multipath propagation (interstellar scattering) that is stronger at low Galactic latitudes.

PALFA Data Acquisition

Data are obtained at Arecibo using the ALFA system, a 7-feed array with cryogenic receivers. Fourteen receiver channels are processed (two polarizations for each feed). The two polarizations are summed before recording to disk.

Current System: Since 2004 PALFA has used correlator systems, the Wideband Arecibo Pulsar Processors (WAPPs) to record correlation functions every $\Delta t = 64\ \mu$s. The RF signal is centered on ~ 1.4 GHz. A 100 MHz bandwidth is mixed to baseband and autocorrelated for 256 lags. Correlation functions are recorded as two-byte integers combined with appropriate header information in a proprietary format. Data set lengths are either 134 s or 268 s, providing 2M or 4M time samples.

The 64 μs sample interval is dictated by the fact that many pulsars have small duty cycles, $W/P_{\rm spin} \ll 1$, (where W is the pulse width), yielding $\sim P_{\rm spin}/W$ harmonics that can be combined into a test statistic (the harmonic sum). To look for spin periods as short as ~ 1 ms and duty cycles of 1/16, we need sampling of 1 ms/16 in order to not quench the very harmonics we would like to sum.

After Fourier transformation, the correlation functions yield 256 spectral channels across 100 MHz, or 0.39 MHz per channel. This channelization allows compensation for dispersive propagation of any pulses from celestial sources. The dispersion delay is

$$\Delta t_{\rm DM} = \frac{4.15\ ms{\rm DM}}{\nu^2}, \quad (1)$$

where ν is the RF in GHz and the dispersion measure to a pulsar at distance D is the integrated electron density n_e,

$$\rm DM = \int_0^D ds\, n_e(s), \quad (2)$$

[1] ALFA = Arecibo L-band Feed Array

expressed in units of pc cm^{-3}. The differential form appropriate for $\Delta\nu \ll \nu$ is

$$\Delta t_{\rm DM} \approx \frac{8.3\ \mu s {\rm DM} \Delta\nu}{\nu^3}, \tag{3}$$

with $\Delta\nu$ in MHz.

"Dedispersion" consists of summing the data over frequency ν after compensating each frequency channel for the dispersion delay. In the blind survey we are conducting, we do not know DM a priori, so we must search over a set of trial values of dispersion measure. Currently our processing searches over $0 \le {\rm DM} \le 10^3$ pc cm^{-3}. The largest value of DM seen so far is ~ 1500 pc cm^{-3} and modeling of the Milky Way's free electron density suggests values as large as 3000. However, these large values obtain for directions not reachable with Arecibo owing to its zenith angle restrictions, e.g. the Galactic center direction. Of course any particular direction at low latitude can have a very large DM if an HII region lies across the line of sight.

For a maximum DM of 10^3 pc cm^{-3}, the total sweep of the pulse across a 100 MHz bandwidth is $8.3\ \mu s \times 10^3 \times 100/1.4^3 \approx$ 300 ms. Removal of dispersion delays can reveal pulses as short as the time resolution, 64 μs.

New System: In early 2008 we will migrate to a new data acquisition system consisting of 14 polyphase filterbank systems that will analyze 300 MHz instead of 100 MHz and with ~ 1024 channels rather than 256. The RF band will be 1.175 to 1.475 GHz. Dwell times per sky position will remain the same (134 s or 268 s, depending on whether we are searching the outer or inner Galaxy).

Data Units: The basic data unit is a *dynamic spectrum* with N_ν frequency channels of bandwidth $\Delta\nu$ that cover a total bandwidth B and N_t time steps separated by Δt that cover a total time T.

Existing data (from WAPPs):

$(B, T) = 100$ MHz $\times$ 268 s

$(\Delta\nu, \Delta t) = 0.39$ MHz $\times$ 64 μs

The recording data rate is 7 signals $\times$ 256 channels $\times$ 2 bytes / 64 μs = 56 MB s^{-1}

New data (starting Q2 2008):

$(B, T) = 300$ MHz $\times$ 268 s

$(\Delta\nu, \Delta t) = 0.30$ MHz $\times$ 64 μs

The recording data rate is 7 signals $\times$ 1024 channels $\times$ m bytes / 64 μs = 112m MB s^{-1}. We are still discussing the value of m. It is advantageous to pack the spectral data into $m = 1$ bytes and possibly into 4 bits.

Data Volumes

The PALFA Consortium is surveying within 5 deg of the Galactic plane for the portion that is reachable with the ≤ 20 deg zenith angle restriction. This requires about 4.7×10^4 separate pointings of the 7-beam system. The total data volume for this survey will reach approximately 1 Pbyte (taking into account that we will reobserve some or all of the positions).

On a yearly basis, we have about 500 hr of telescope time for the PALFA survey. Some of this (about 10%) is used for initial follow-up timing of discovered pulsars. Overhead (telescope slewing, calibration) consumes another 10%. So of order 400 hr of actual survey data are obtained.

The total annual data volume is

WAPPS: 56 MB/s $\times$ 400 hr $\times$ 3600 s/hr = 80.6 TB/yr

New spectrometers: 112m MB/s $\times$ 400 hr $\times$ 3600 s/hr = 161 TB/yr

A data set from a single telescope pointing consists of the spectrometer data for 7 telescope beams. For our typical 268 s data sets with 64 μs sampling, the data sets are 4M samples long, corresponding to 15 GB per data set with the WAPPs and double this with the new spectrometers for $m = 1$ byte/sample.

PALFA Data Analysis

Pulsars that are relatively steady in amplitude[2] are found by a search over period P, DM duty cycle= (pulse width)/period, which determines the number of detectable harmonics in the power spectrum), and acceleration. The acceleration analysis is appropriate for binaries with periods $T \lesssim 0.1 P_{\rm orb}$ for which the phase perturbation from the orbit can be approximated as a parabola.

The processing pipeline consists of

[2] As opposed to detecting some objects that emit isolated bursts, such as the RRATs.

1. Unpacking of two byte integer data and short FFTs to convert correlation functions to spectra; this amounts to 4M FFTs of length 256 with WAPP data; the new spectrometer produces spectra directly so this step will not be needed.
2. Dedispersion with $N_{\rm DM}$ trial values of DM (currently $N_{\rm DM} = 1270$)
3. Search for isolated pulses in the time series for each DM (matched filter analysis and cluster analysis)
4. Search for periodic signals in each time series through a long FFT + harmonic sum analysis
5. For some processing sites, we also search vs. acceleration with the goal of processing all the data along the acceleration axis as well as the other axes.

The basic post-dedispersion analysis unit is therefore a *time series* of length $T/\Delta t = 268\ s/64\ \mu s = 4$Ms. Our existing pipelines analyze these units and meta-analyze aggregate results on the many time series for a given telescope pointing (i.e. all trial DMs for the 7 ALFA beams). Further meta-analysis is done on the candidate lists from different pointings to filter out candidates from RFI.

At Cornell, we process a 7-beam data set in about 2 hr with one node on a Unisys Itanium cluster (c. 2004) devoted to each beam. At the guesstimate level, about 20% of this is for unpacking and other overhead, 20% for dedispersion, 15% for single-pulse detection, 30% for long FFTs, and 15% for harmonic summing, candidate identification and synchronous averaging at candidate pulse periods.

The amount of processing time for a single time series dedispersed with a particular trial DM is $\sim 0.45 \times 2.5$ hr/1271 = 3.2 s. With current processors this is probably more like 1 s.

Pulse Phase Perturbations from Orbits

Consider just a single harmonic of the Fourier series that describes a strictly periodic pulse train. Sharp pulses produce harmonics up to $\ell_{\rm max} = 32$ or so, so degradation of the amplitude of high-order harmonics should be considered. For weak pulsars, it is reasonable to consider $\ell_{\rm max} = 8$. Now consider the ℓ-th harmonic. The strictly periodic phase of this harmonic in the rest frame of the pulsar is modified by orbital motion. The variable part of the pulsar's distance along the line of sight is

$$z(t) = a_1 \sin i \sin(\Omega t + \Phi_0) \equiv A \sin(\Omega t + \Phi_0) \tag{4}$$

for a circular orbit of radius a_1, inclination i, orbital frequency $\Omega = 2\pi/P_{\rm orb}$, and orbital phase Φ_0. Ignoring the propagation time across the pulsar-Earth distance and assuming non-relativistic pulsar motion, the arrival time is $\tau = t + z(t)/c$ and the Fourier phase is $\phi(t) = f_{\rm spin}\tau$ (cycles) with $f_{\rm spin} = 1/P_{\rm spin}$.

Pulsars spin down from magnetic torques so f decreases with time (apart from spin-ups in rare glitches). Practically, however, in a pulsar search data set of a few hundred seconds, there is negligible contribution to the phase from $\dot f_{\rm spin}$ and $\ddot f_{\rm spin}$ from spindown. Pulsar motion in compact binary systems contributes significantly for orbital periods of 10 hr or less. In the following $f_{\rm spin}$ is assumed constant. If there should be an object with very high spindown rate, the contributions to the derivatives of $f_{\rm spin}$ will be combined with orbital contributions.

Expanding the Fourier phase we have

$$\phi_\ell(t) = \phi_0 + f_\ell t + \frac{1}{2}\dot f_\ell t^2 + \frac{1}{6}\ddot f_\ell t^3, \tag{5}$$

where, for general $z(t)$,

$$f_\ell = \ell f_{\rm spin}\left[1 + c^{-1}\dot z(0)\right] \tag{6}$$

$$\dot f_\ell = \ell f_{\rm spin} c^{-1}\ddot z(0) \tag{7}$$

$$\ddot f_\ell = \ell f_{\rm spin} c^{-1}\dddot z(0). \tag{8}$$

For a circular orbit

$$a_1 = 0.5 R_\odot \frac{[M P_{\rm orb}^2({\rm hr})]^{1/3}}{(1 + M_1/M_2)} = 1.17\,{\rm sec}\,\frac{[M P_{\rm orb}^2({\rm hr})]^{1/3}}{(1 + M_1/M_2)}, \tag{9}$$

where $M \equiv M_1 + M_2$ is the total mass (in units of $M_\odot$) and M_1 is the pulsar mass. Maximum values for velocity, acceleration and jerk are

$$v_{\rm max} = \dot z_{\rm max} = a_1 \sin i\ \Omega = 611\ {\rm km\ s^{-1}}\ \left[\frac{1}{P_{\rm orb}({\rm hr})}\right] \frac{[M P_{\rm orb}^2({\rm hr})]^{1/3}\sin i}{(1 + M_1/M_2)} \tag{10}$$

$$a_{\rm max} = \ddot z_{\rm max} = a_1 \sin i\ \Omega^2 = 10^{5.0}\ {\rm cm\ s^{-2}}\ \left[\frac{1}{P_{\rm orb}({\rm hr})}\right]^2 \frac{[M P_{\rm orb}^2({\rm hr})]^{1/3}\sin i}{(1 + M_1/M_2)} \tag{11}$$

$$\dot a_{\rm max} = \dddot z_{\rm max} = a_1 \sin i\ \Omega^3 = 10^{2.3}\ {\rm cm\ s^{-3}}\ \left[\frac{1}{P_{\rm orb}({\rm hr})}\right]^3 \frac{[M P_{\rm orb}^2({\rm hr})]^{1/3}\sin i}{(1 + M_1/M_2)} \tag{12}$$

The maximum values of acceleration and jerk correspond to maximum values for the pulse-frequency derivatives of

$$\dot f_{\ell,\max} = \frac{\ell f_{\rm spin} a_{\max}}{c} = \frac{\ell f_{\rm spin} a_1 \sin i}{c}\left(\frac{2\pi}{P_{\rm orb}}\right)^2 = 10^{-2.5}\ {\rm Hz\ s^{-1}}\ \frac{\left[M P_{\rm orb}^2({\rm hr})\right]^{1/3}\sin i}{1+M_1/M_2}\left[\frac{\ell}{P_{\rm spin}({\rm ms})}\right]\left[\frac{1}{P_{\rm orb}({\rm hr})}\right]^2, \quad (13)$$

$$\ddot f_{\ell,\max} = \frac{\ell f_{\rm spin} \dot a_{\max}}{c} = \frac{\ell f_{\rm spin} a_1 \sin i}{c}\left(\frac{2\pi}{P_{\rm orb}}\right)^3 = 10^{-5.2}\ {\rm Hz\ s^{-2}}\ \frac{\left[M P_{\rm orb}^2({\rm hr})\right]^{1/3}\sin i}{1+M_1/M_2}\left[\frac{\ell}{P_{\rm spin}({\rm ms})}\right]\left[\frac{1}{P_{\rm orb}({\rm hr})}\right]^3. \quad (14)$$

Merger Time Scale and Shortest $P_{\rm orb}$ to Search For

Taking only gravitational radiation into account, the time to merge for an initial orbital radius a_0 (circular orbit) is

$$T_0 = \frac{5}{256}\frac{c^5 a_0^4}{G^3 M_1 M_2 M} = 7.1\ {\rm Myr}\ P_{\rm orb}^{8/3}({\rm hr})\left(\frac{1.4\ M_\odot}{M_1}\right)^{5/3}\left[\frac{(1+M_2/M_1)^{1/3}}{2^{1/3}M_2/M_1}\right] \quad (15)$$

(defined so that in the rightmost equation, the mass factors are unity for a NS-NS binary with $M_1 = M_2 = 1.4\ M_\odot$). The merger time for an eccentric binary will be substantially shorter.

Shortest Expected $P_{\rm orb}$: A rule of thumb for the smallest $P_{\rm orb}$ to search for follows from consideration of the merger rate and merger time scale. The merger rate for NS-NS binaries in the Galaxy is approximately $\dot N_{\rm merge} \approx 10^{1.4\pm0.5}\ {\rm Myr}^{-1}$. Setting to unity the product of the merger rate and merger time, $\dot N_{\rm merge} T_0$, yields the orbital period at which there will be one system in the Milky Way (in the mean) with a remaining lifetime $\dot N_{\rm merge}^{-1}$. The numbers yield an orbital period of about 6 to 13 min with a lifetime before merger of $10^{4.1}$ yr to $10^{5.1}$ yr. For data sets of length ~ 5 min, searching down to 6 min orbital periods requires a full three-parameter search rather than an acceleration + jerk search.

For a search that covers a wide range of orbital periods, some with $P_{\rm orb} \sim T$ and most with $P_{\rm orb} \gg T$ (e.g. searching to 8 hr periods with $T = 5$ min), the search algorithm might be optimized by using a circular-orbit parameterization for short periods and a cubic polynomial for long periods. If so, then how does the transition region go?

Fourier Analysis of Time Series Corrected for Acceleration and Jerk

Suppose we analyze a dedispersed time series by first regridding the time series using values of acceleration and jerk followed by a Fourier analysis and harmonic summing. The Fourier analysis "finds" the value of f_ℓ to within $\delta f \approx T^{-1}$, implying an error of $\ell/2$ cycles at $\pm T/2$. With high S/N, the frequency can be identified to better than this. Values for acceleration and jerk are chosen from a uniform grid of values with spacings $\delta \dot f_\ell$ and $\delta \ddot f_\ell$. The error in phase associated with a maximum error of 1/2 grid spacing from either of the true values is

$$\delta\phi_\ell(t) = \frac{1}{2}\left[\frac{1}{2}\delta\dot f_\ell t^2 + \frac{1}{6}\delta\ddot f_\ell t^3\right]. \quad (16)$$

Over a time series of length T the maximum error at $\pm T/2$ can be required to be less than ϵ cycles. Typically, pulses have duty cycles $\lesssim 10\%$ so we do not want to tolerate more pulse smearing than about this amount; otherwise S/N is reduced. This implies constraints on the grid spacings

$$\delta\dot f_\ell \le \frac{16\epsilon}{T^2} \quad (17)$$

$$\delta\ddot f_\ell \le \frac{96\epsilon}{T^3}. \quad (18)$$

The implied total number of values in each grid dimension is then

$$N_{\dot f} = \frac{2\dot f_{\ell,\max}}{\delta\dot f_\ell} = \frac{\ell f_{\rm spin} a_1 \sin i}{2\epsilon c}\left(\frac{\pi T}{P_{\rm orb}}\right)^2 \approx \frac{401}{\epsilon_{0.1}}\frac{\left[M P_{\rm orb}^2({\rm hr})\right]^{1/3}\sin i}{1+M_1/M_2}\left[\frac{\ell}{P_{\rm spin}({\rm ms})}\right]\left[\frac{T/(5\ {\rm min})}{P_{\rm orb}({\rm hr})}\right]^2 \quad (19)$$

$$N_{\ddot f} = \frac{2\ddot f_{\ell,\max}}{\delta\ddot f_\ell} = \frac{\ell f_{\rm spin} a_1 \sin i}{6\epsilon c}\left(\frac{\pi T}{P_{\rm orb}}\right)^3 \approx \frac{35}{\epsilon_{0.1}}\frac{\left[M P_{\rm orb}^2({\rm hr})\right]^{1/3}\sin i}{1+M_1/M_2}\left[\frac{\ell}{P_{\rm spin}({\rm ms})}\right]\left[\frac{T/(5\ {\rm min})}{P_{\rm orb}({\rm hr})}\right]^3. \quad (20)$$

Here we have used $\epsilon \equiv 0.1\epsilon_{0.1}$. What harmonic number ℓ should we consider? We want no more than about 10% of pulse phase error in order that the pulse not be smeared by more than this amount. An error in *harmonic* phase of ϵ corresponds to an error in pulse phase of ϵ/ℓ. Given that we expect harmonic sums to maximize for sums of 8 to 16 harmonics, it is reasonable to just consider a 10% error on the fundamental frequency, $\ell = 1$.

Comments:

1. Nominal coefficient values in Eq. 19-20 are for a spin period of 1 ms. Pulsars with NS companions tend to have periods of 20 ms or longer, so if we target these objects and pulsars with BH companions, the grid size becomes substantially smaller.
2. Other factors can increase the grid size, including the total mass M (though slowly as the 1/3 power) and the orbital period. Sub-hour orbits are of great interest, even for WD companions. However pulsars with WD companions tend to have millisecond periods, so the grid size could be very large.
3. One option, not particularly desireable, is to process data in smaller pieces than the length $T = 268$ s.
4. A reasonable option is to choose an astrophysically selected subset of the overall phase space to explore in order to match the processing requirements to network and processing capacity.
5. The number of grid values $N_{\dot{f}} \times N_{\ddot{f}}$ may be less than given in the above equations owing to likely redundancies in particular pairs of values yielding effectively the same phase function $\delta\phi_\ell$ over the interval of interest.

Possible Processing Regimes

NS-NS Binaries: The most interesting binaries — and the most difficult to detect — are those with the shortest periods. These typically have high eccentricities, so the accelerations and jerks at periastron will be much larger, for a given period $P_{\rm orb}$, than as calculated for circular orbits. With $M = 2.8M_\odot$, $1 + M_1/M_2 = 2$ and $P_{\rm spin} \geq 20$ ms we have, for circular orbits, the number of grid points to search is the product

$$N_{\dot{f}} N_{\ddot{f}} \approx 17 P_{\rm orb}({\rm hr})^{-11/3} \left(\frac{20\ {\rm ms}}{P_{\rm spin}}\right)^2. \tag{21}$$

Suppose we target orbital periods as short as 0.5 hr; the coefficent becomes 200; multiplying by a factor of 10 to 100 increases it to 2000 to 20,000. With ~ 1 s of processing for each $\dot{f}, \ddot{f}$ pair, the latter case implies that it will take 1/4 day of processing per time series. With 1270 DM values and 7 ALFA beams, it takes $\sim 10^4$ processors 1/4 day to process a single pointing of 5 min.

NS-WD Binaries: These binaries typically consist of a millisecond pulsar (MSP) in a circular orbit with a long orbital period. The shortest period MSP-WD binary is 6 hr (J0751+18). All the known MSP-WD binaries have very long merger times, suggesting that most will have periods longer than 6 hr. However, short-period MSP binaries have been highly selected against in previous surveys, biasing our view, and we should search down to $P_{\rm orb} = 1$ hr, say. Although the fastest MSP has $P_{\rm spin} = 1.4$ s, most have periods longer than 3 ms, so that could be used as a basis for reducing processing requirements. Per time series, the number of grid points to search for $P_{\rm spin} \geq 3$ ms is

$$N_{\dot{f}} N_{\ddot{f}} \approx 166 P_{\rm orb}({\rm hr})^{-11/3} \left(\frac{3\ {\rm ms}}{P_{\rm spin}}\right)^2, \tag{22}$$

where we have used a WD mass of $M_2 = 0.5M_\odot$.

There is another possibility for limiting the search space for MSPs. Multipath propagation in the interstellar medium (ISM) causes uncorrectable pulse smearing that becomes very severe for $DM \gtrsim 500$ pc cm^{-3} at 1.4 GHz. So we could limit the acceleration + jerk search to DM less than this amount.

Better: Searching for Circular Orbits

If the goal is find objects with orbital periods as small as the data-span length e.g. $P_{\rm orb} \sim T$, then a direct search over the parameters A, Ω, Φ_0 of Eq. 4 of a circular orbit is needed rather than a Taylor-series approach. There is probably considerable degeneracy in the parameter space given that most orbital periods of interest (e.g. 10 min to 8 hr) will be much longer than the data set length, $T \ll P_{\rm orb}$, for which an acceleration + jerk search is ok. Approximate step sizes in the grid for a phase error ϵ are (where we treat A as having time units):

$$\delta A \sim \frac{\epsilon P_{\rm spin}}{\min[\Omega T, 1]} \tag{23}$$

$$\delta\Omega \sim \frac{2\epsilon P_{\rm spin}}{AT} \tag{24}$$

$$\delta\Phi_0 \sim \frac{\epsilon P_{\rm spin}}{A}. \tag{25}$$

Action item: we need to figure out the minimal grid to cover orbits of interest. We can take into account that the algorithm consists of regridding of the time series followed by the FFT and harmonic summing phases. The mean doppler shift from the piece of the orbit sampled is absorbed by the Fourier analysis. I.e. the regridding need only take out quadratic and higher terms and can place the resulting harmonics into the wrong Fourier bins with no S/N penalty. It will still be detected and the true frequency can be sorted out after the fact. Roughly, the mean Doppler shift is

$$\frac{\Delta f}{f_{\rm spin}} \sim \frac{[z(T) - z(0)]}{cT}. \tag{26}$$

So, the function to use for rebinning the time series is

$$\Delta z(t) = z(t) - \frac{[z(T) - z(0)]\, t}{T}. \tag{27}$$

The number of grid values in the parameter 3-space of $\Delta z(t)$ is large and needs to be kept to a minimum for a fixed S/N loss from gridding. Thus $\Delta z(t)$ needs to be analyzed over reasonable ranges of A, Ω and Φ_0 to find a minimal grid.

To define the grid, we can build in some astrophysics if we have to. For example, a NS-NS binary with $P_{\rm orb} = 6\ hr$ today will have a detected, recycled pulsar with $P_{\rm spin} = 50$ ms that becomes, say, 500 ms when the orbital period is 0.5 hr. The grid can be more coarse for such compact binaries with longish spin periods. But maybe we don't want to build in too much convention. Perhaps nature knows a way to produce MSPs in sub-hour orbits!

Numbers of Trials and Significance Levels

For PALFA's non-acceleration searches, we use a threshold of about 8σ in the harmonic sum analysis. Searching over $N_{\rm grid}$ parameter sets implies a proportional increase in number of trials. The threshold must increase accordingly to keep the same false-alarm rate. This might raise the threshold to 20σ or more, but that does not preclude detection of new pulsars. Many of PALFA's detections of non-binary pulsars are at the $\gg 10\sigma$ level. Moreover there could be very strong pulsars that have been missed in periodicity surveys because they are in tight binaries. Some such objects may even have been catalogued in continuum surveys such as the NVSS (NRAO VLA Sky Survey).

Grid Computing with Einstein@Home Resources

As we discussed in Boston in Dec 2007, a reasonable data unit for processing on an Einstein@Home client is a single dedispersed time series, of 4M samples. Current processing at Cornell produces these as floats but with scaling we could use single-byte integers. The signals we are looking for are buried in the noise so 1 byte should be sufficient. Most time series are afflicted by RFI and occasional large pulses from known, strong pulsars. We can clip non-noise events as part of the scaling process. The Cornell code does a non-acceleration analysis and has nearly kept up with the data flow. PALFA Consortium members (particularly UBC, McGill and UT Brownsville) run pipelines that include an acceleration analysis, which takes substantially longer and has not kept up with the analysis. However, processing capacity is coming on line at the TACC that will fill the gap on acceleration searches. Adding even a jerk parameter to the analysis is prohibitive at this point for Consortium sites.

A value-added analysis for the Einstein@Home community is a search for pulsars in compact binaries, either through a two-parameter search (acceleration + jerk) or a three-parameter search over circular orbits (or some hybrid). The scaling laws given above should allow us to identify a meaningful subspace that can be processed to an interesting level of completeness and with the constraint that the processing time per data unit is approximately one day or longer.