# Archiving the 350-MHz GBT drift-scan survey

D.R. Lorimer (WVU), S.M. Ransom (NRAO) & M.A. McLaughlin (WVU)

We propose to purchase 80 TB of portable disk space and a server workstation which would allow us to transport and archive data for our upcoming GBT drift-scan survey at 350 MHz. The primary science goal of the survey is to discover millisecond pulsars that have a rich variety of applications, including tests of general relativity, probes of stellar evolution and sensitive detectors of gravitational waves in the nHz regime. Once we have analyzed the data at remote processing sites, the disks would be shipped back to NRAO where they would form a permanent archive, available to anyone wanting to investigate the 350-MHz radio sky.

## A 350-MHz GBT drift scan survey

Large-scale Galactic surveys have uncovered more than 1700 pulsars over the past 40 years. Among this population are many jewels, including ∼60 clock-like millisecond pulsars (MSPs; Backer et al. 1982, Nature, 300, 615), dozens of interesting binary systems (e.g. the double-pulsar: Burgay et al. 2003, Nature, 426, 531; Lyne et al. 2004, Science, 303, 1153) and a handful with exotic behavior such as free precession (Stairs et al. 2000, Nature, 406, 484), periodically active emission (Kramer et al. 2006, Science, 312, 549), and mysterious transient radio bursts (McLaughlin et al. 2006, Nature, 439, 817).

We are currently in a golden age of pulsar surveys. New telescopes and data acquisition systems are opening up new areas of parameter space that were previously inaccessible. As part of this work, we have been granted 900 hrs to utilize the summer shutdown period of the GBT. Observing in a drift-scan mode at 350 MHz, we expect to cover about 20% of the entire sky (2.6 steradians). The data we collect will be investigated by two main groups: approximately 700 hours will be analyzed by a collaboration of astronomers and graduate students from WVU, NRAO, Cornell, UBC, Berkeley and Amsterdam. The remaining 200 hours will be distributed to schools in West Virginia for analysis by high-school students as part of the 'Pulsar Search Collaboratory' — a novel project which aims to improve the IT skills in schools and give students first-hand experience in a real research project.

For both projects, we will be using the PF1 340 MHz receiver on the GBT centered at 350 MHz along with the SPIGOT in a 50-MHz bandwidth mode, providing 2048 8-bit lags of summed polarizations sampled every 81.92 $\mu$s. The excellent resolution ensures that we are sensitive to the shortest period MSPs with dispersion measures up to $\sim$100 cm$^{-3}$ pc. Given the pointing constraints of the telescope, we will observe primarily sky in the declination range $-40° < \delta < +40°$. Our rate of sky coverage will be approximately 1 steradian every 17 days.

The main aim of these observations is to find new MSPs. In addition to furthering our knowledge about their Galactic population, new discoveries also spawn a number of follow-up studies

including tests of general relativity, white dwarf cooling and binary evolution and the direct detection of gravitational radiation. To estimate the likely MSP yield, we have carried out simulations assuming MSP population properties similar to those inferred by Cordes & Chernoff (1997 ApJ, 482, 971). For an all-GBT-sky 350-MHz drift survey, we find an isotropically detectable sample of objects with a surface density of *10 detectable MSPs per steradian.* Taking into account the 40 MSPs currently known in the GBT sky, we predict about 10–15 MSPs will be discovered by the survey. Similar predictions for the normal pulsar population suggest an additional 50–100 normal pulsars will be discovered.

## Archiving the survey data

We expect the 350-MHz survey to be a unique probe of the 350-MHz radio sky. In addition to finding new and exciting pulsars, the data will also be mined for transient radio signals such as those discovered in the Parkes survey by McLaughlin et al. (2006) and could possibly be used to produce a 350-MHz continuum map. Based on our previous experience with the Parkes multibeam survey, which is still being mined with newly developed algorithms and producing new discoveries, a permanent archive will be a fantastic resource. Since a survey of this scale may not be possible to repeat for many years to come, we believe that it is very important to archive it at NRAO as a resource for the astronomy community in general. For around 900 hours of observing, we expect to collect 80 TB during the survey. Given our current disk-space resources, which are in demand from other projects, it is not feasible for us to keep the data for more than about a year after collection.

To create a permanent archive, we propose to purchase a dedicated set of portable disks using funds from the Lockheed Martin remediation. For an 80 TB data set, the survey could be stored on $110 \times 750$ GB disks. Based on current market prices, the most cost effective disks appear to the Seagate Pushbutton dual interface drives which retail at $320 each. We also request $2k for a PC server which could be used initially for intermediate processing at the Green Bank site and, later, to serve the disks after they become publicly available. The total estimated cost would therefore be $37k. This solution has the added benefit that the disks can be plugged into any machine with a FireWire or USB2 interface. They could therefore be very easily used to transfer the data off the SPIGOT. Getting data off of the SPIGOT quickly is absolutely essential for us to make optimal use of the summer shutdown observing. Such disks are also optimal for shipping to remote sites and onto short-term storage for rapid data reduction.

Given the resources currently available to us. We expect to complete our processing by January 2009, and have all the disks shipped back to NRAO for long-term storage. This coincides with the 18-month NRAO proprietary period so that the raw data would then be available to anybody wishing to access part of the survey, or make a copy of the entire archive. A searchable database will be developed during the survey which will allow clear identification of the appropriate datafiles which cover a requested sky region. Software to read the raw data and do basic tasks such as create spectra and do time series analysis will be publicly available.