THE ASTROPHYSICAL JOURNAL, 366:L33–L37, 1991 January 1

REFRACTIVE INTERSTELLAR SCINTILLATION AND MILLISECOND PULSAR TIMING

WAYNE HU[1]
Princeton University
ROGER W. ROMANI[2]
Institute for Advanced Study, School of Natural Science, Princeton, NJ 08540
AND
DANIEL R. STINEBRING[3]
Princeton University
Received 1990 August 10; accepted 1990 October 16

ABSTRACT

Timing observations of the millisecond pulsar PSR 1937+21 now show significant low-frequency noise. This noise is incompletely removed by a simple dispersion-correcting algorithm based on dual frequency observations. We have developed a two-dimensional simulation of refractive interstellar scintillation (RISS) propagation effects, with parameters appropriate to the pulsar timing experiment. These computations allow us to quantitatively assess agreement with the interstellar model and show good agreement with a power-law spectrum of interstellar tubulence. Moreover, a second roughly frequency-independent noise component is isolated from the data. This noise, with a very steep red spectrum, may be best explained as irregularities in the pulsar rotation or, possibly, as signature of a cosmological gravitational wave background.

Subject headings: interstellar: matter — pulsars — stars: individual (PSR 1937+21)

1. INTRODUCTION

Millisecond pulsar timing has become a powerful cosmological probe. In particular, departures of the arrival time data from a simple model of the pulsar spindown and position are so small that significant constraints can be placed on any unmodeled effect, such as a cosmological background of ultra–low-frequency gravitational waves. Recently Stinebring et al. (1990, hereafter SRTR) have applied such an analysis to 7.1 yr of high-quality timing observations of the original millisecond pulsar, PSR 1937+21. In these data, significant departures from the simple spindown model are appearing on time scales comparable to the observation span. This noise source is evidently very "red," i.e., it has a power spectrum whose amplitude rises steeply to lower frequencies. In identifying the source of this noise, it is useful to measure its spectral index. Interesting candidate noises include interstellar propagation effects, inaccuracies in Earth time and ephemeris scales, instabilities in the pulsar rotation period, and metric perturbations due to a stochastic gravity wave background. All of these have red spectra with a range of power-law indices. Due to bandpass leakage, Fourier transform techniques are completely inadequate as power estimators for such spectra. Accordingly, SRTR have employed a method based on orthonormal polynomial fitting that provides a good measure of the red noise spectral content, a useful discriminant between different putative sources.

One source of perturbation to the arrival time, certainly present at a significant level, is due to variations in the intervening column of interstellar electrons caused by motion of the Earth-pulsar line of sight (Armstrong 1984; Blandford, Narayan, & Romani 1984). In order to monitor these fluctuations, observations have been made of PSR 1937+21 at two radio frequencies since 1984. In principle, the dispersion measure (DM $= \int n_e\, ds$) variations in the intervening electron column can be corrected for by removing a term proportional to $\lambda_{\rm obs}^2$ from the residuals to the timing model (Rawley, Taylor, & Davis 1988; see also Cordes et al. 1990). However, two peculiarities with the resulting arrival time residuals were noticed after implementation of this correction scheme (SRTR). First, the δDM correction did not fully remove the low-frequency noise in the data. Second, the noise power at longest time scales in the 2380 MHz data actually increased somewhat.

These results suggest two interpretations: either our algorithm for removing interstellar propagation effects is inadequate, or there is an additional source of low-frequency noise in the data. Discriminating between these is important since, insofar as timing noise can be attributed to known sources, bounds on other effects, such as gravity wave perturbations, will improve. Further, these dual frequency observations constitute some of our best measurements of interstellar electron density fluctuations, and it is important that the application be well understood. The dominant component of the interstellar perturbations is believed to be the total electron column (δDM) variations which give rise to arrival time residuals proportional to λ^2. However, there are angle of arrival and other effects with different frequency dependencies (e.g., λ^4) which also may be present. The interstellar electron density fluctuations apparently have a nearly power-law spectrum over a substantial range of wave scales (Armstrong, Cordes, & Rickett 1981). At spatial frequencies q of interest, the spectrum is believed to follow the Kolmogorov power spectrum, $Q(q) \sim q^{-11/3}$. Analytic estimates of the timing residuals expected from such spectra (Romani, Narayan, & Blandford 1986; Cordes, Pidwerbetsky, & Lovelace 1986) indicate that

[1] Postal address: Department of Physics, University of California, Berkeley, Berkeley, CA 94720.
[2] Also: Department of Physics, Stanford University, Stanford, CA 94305-4060.
[3] Postal address: Department of Physics, Oberlin College, Oberlin, OH 44074.

the "refractive" angle of arrival contributions should be somewhat smaller than the residual noise seen. However, the results of such refractive interstellar scintillation (RISS) computations apply to the mean, ensemble average response to a power spectrum. A particular realization of a red noise process, such as our arrival time series, may behave quite differently than the ensemble mean. Foster & Cordes (1990, hereafter FC) have employed a numerical simulation of RISS-induced perturbations to the pulsar arrival time to examine particular electron density fluctuations and to assess the importance of the various terms in the propagation model.

To improve our understanding of the dual frequency data, we have developed a similar numerical simulation of the refractive interstellar medium, extending the analysis to a two-dimensional phase screen and concentrating on parameters appropriate to the present PSR 1937+21 timing experiment. Analyzing many realizations of these simulations with the same techniques applied to the real data by SRTR, we can calibrate, in a Monte Carlo sense, the response of our arrival time observables to the scintillations. Most importantly, this computation allows us to assign probabilities for a given degree of departure from the expected behavior, and hence quantify the significance of the non-λ^2 terms above. We describe these simulations and the analysis of PSR 1937+21 timing data in the next section. In the conclusion, an assessment of the agreement with the standard model for the interstellar electron turbulence is made, along with implication of these results for other noise sources, such as gravity wave perturbations.

2. SIMULATION AND DATA ANALYSIS

We replace the interstellar medium along the distance D_p to the pulsar by an equivalent thin phase-changing screen at a distance D_s (cf. Romani et al. 1986; Cordes et al. 1986). Interstellar scattering is generally quite strong, and there are many orders of magnitude between the refractive scale r_r characterizing the total scatter-broadened image size and the diffractive scale r_d, which gives the distance on the screen over which the phase change is $\sim\pi$. The three-dimensional spectrum of electron density fluctuations is given by $\Phi(q) = Q_0 q^{-\beta}$, with $q \sim 1/r$ and β near the Kolmogorov value, 11/3. With such a power-law spectrum we can relate (Romani et al. 1986) the refractive and diffractive length scales by

$$r_r r_d = \lambda D_s \left[\frac{\beta - 2}{4(4-\beta)}\right]^{1/2} . \quad (1)$$

Moreover, the decorrelation bandwidth, the frequency over which the intensity on the ground decorrelates by $1/e$, can be related (FC, eqs. [4.3]–[4.11]) to the diffractive length scale by

$$r_d^2 = \pi^{(6-\beta)/(\beta-2)} \frac{D_s(D_p - D_s)\Delta\nu_d \lambda^2}{2cD_p} , \quad (2)$$

where $\Delta\nu_d$ is the decorrelation bandwidth at observation wavelength λ. This observable allows us, in principle, to normalize our simulation; inconsistencies in the definitions of $\Delta\nu_d$ and r_d leave small error factors of order unity.

Variations in the line-of-sight interstellar medium have two principal effects on the pulsar arrival time (Blandford & Narayan 1985). The first, a dispersive delay caused by the excess electron column, is

$$\delta t = \frac{\lambda\Delta\phi}{2\pi c} = \frac{r_e \lambda^2}{2\pi c} \int_0^{D_p} \delta n . \quad (3)$$

The second is geometric due to refraction-induced angle-of-arrival changes

$$\delta\theta = \frac{-\lambda}{2\pi}\left[\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2\right]^{1/2} . \quad (4)$$

This effect is manifest in two ways: as a simple path length variation in the time of flight ($\sim\theta^2 = \lambda^4$), and as a perturbation to the apparent position and proper motion on the sky not present in the timing model (principally, this appears as a position fit error giving residuals at a period ~ 1 yr, called the ISS barycentric correction).

The study by Foster & Cordes (1990) has attempted to model RISS arrival time perturbations using two one-dimensional phase screens. However, this approximation neglects correlations removed from, but parallel to, the pulsar line of travel. We therefore extend the analysis to a full two-dimensional phase changing screen placed at $D_s = 1/2D_p$, half the distance to the pulsar. We construct the two-dimensional random phase screens in wavenumber space, scaling the frequency components by a power law, $\propto f^{-\beta/2}$. This is then transformed to real space and normalized such that the squared first differences of the phases, the phase structure function $D_\phi(b) = \pi^2(b/r_d)^{\beta-2}$, has the appropriate value at the grid spacing of the simulation, b.

In constructing simulations appropriate to our time series for PSR 1937+21, we connect the linear dimension of the screen to the observation span through the relative velocity between the Earth-pulsar line and the density perturbations. PSR 1937+21 appears to be a low-velocity object (Rawley et al. 1990), so the dominant term in this relative velocity will be shear due to differential Galactic rotation along the line of sight; we take $v_{\rm rel} = 50$ km s^{-1}. We are principally interested in the long-wavelength, low-frequency perturbations so our simulation must extend somewhat farther than the line-of-sight travel during the total observation span, 7.1 yr. Further, to allow calibration of the dispersion measure corrections, we simulate both the 1.4 GHz and 2.4 GHz observations. At the higher frequency, the refractive scattering disk will be smaller (by λ^2; cf. eqs. [1] and [2]), so to compute accurately the phase gradients and bending angles across the pulsar image we require several grid points across r_r at 2.4 GHz. Computational expediency requires that the two-dimensional screen be symmetric, but we find that the dynamic range constraints above are adequately satisfied with a 1024 × 1024 element grid. This results in a 5 day, 2.1×10^{12} cm grid spacing.

In the computation, incident rays are passed through the phase screen, traced to the observer plane, and recorded on the matching ground grid. In addition to the intensity of the incident ray, this grid contains information on the excess phase accumulated and the position of origin on the phase screen. This allows computation of the arrival time perturbation (as well as its dispersive and geometric components) along with the intensity pattern on the observer plane. Taking this plane as our frame of reference, we follow the traveling pulsar line of sight, monitoring the refractive variations recorded in the computations. Since unmodeled diffractive processes scatter the refracted rays through a range of angles at the screen, we convolve the observer plane along this path with a smoothing function corresponding to the scatter-broadend beam. For a spectrum with $\beta = 4$ this beam will be a two-dimensional Gaussian; for the $\beta = 11/3$ case, the small departures from the Gaussian model were found to be negligible and the Gaussian form was adopted for considerable reduction in computation

time. After this convolution, the time series of the RISS-induced arrival time perturbations constitute the principal output of the simulation. With the same phase screen, renormalizing and rescaling the bending angles and scattering disk to correspond to 2.4 GHz, we construct a second time series to complete the dual frequency data.

We sample the resulting perturbation time series at the same epochs as the observed data for PSR 1937+21 and treat these "arrival times" exactly as the Arecibo data set. To produce a set of timing residuals, we pass the artificial data through the standard pulsar timing package TEMPO (Taylor & Weisberg 1989). We may then apply any desired analysis scheme to this residual set and compare with the response of the physical data. This is a second important feature of our study; we have generated a large number of random screens and therefore can use the full distribution of our selected observables for our modeled data to allow a quantitative assessment of the physical data's agreement with the fiducial model. Note that a number of independent pulsar tracks can be accommodated on a single realization of the phase changing screen. Spacing tracks by 6 times the Gaussian width at 1.4 GHz, we obtained 5600 independent realizations from 200 computed screens.

We compared our results with the data obtained at 1408 MHz and 2380 MHz over 7.1 yr with the Arecibo telecope. Details of the instrumentation and observations as well as further information on the reduction procedures can be found in Rawley et al. (1988) and SRTR. In addition to the two arrival time data sets, a time series of the variation in the mean dispersion measure to the pulsar was obtained over 5.2 yr by subtracting a smoothed ($\sim 50^d$ Gaussian running mean) version of the 1.4 GHz residuals, rescaled by λ^2, from the 2.4 GHz data. Such δDM time series were also derived for the artificial data. The observables used in comparing these artificial and real time series are the amplitudes of the noise spectral power in several octave bands at the lowest accessible frequency of interest. As noted above, we have employed a power measure based on the amplitude of fitted orthonormal polynomials. In brief, we fitted four polynomial functions, orthonormal over the observed data epochs, to the entire observed (or simulated) time series. The first three polynomials are degenerate with fitted parameters of the pulsar spindown model and contain no information. The squared amplitude of the fourth, cubic, polynomial comprises our power spectral estimator on the lowest accessible frequency range—a cubic appears roughly as a sinusoid with one cycle over the observation span. The data are then divided into $m = 2, 4, 8$ pieces, the fourth orthonormal coefficient is estimated for the m subsets, averaged, and a power spectrum with octave sampling is thus formed. Details and analysis of the method are found in SRTR.

We first consider the time series of the DM corrections in the simulated and the real data. Examination of the several RISS-induced components of the simulated arrival time delay show their relative amplitudes to be in agreement with analytic results. At both frequencies, the dispersive δDM component completely dominates for noise wavelengths $\gtrsim 1$ yr; at 2.4 GHz, the "barycentric correction" term becomes significant for somewhat shorter noise periods, while the geometric time-of-flight term is always small. Using the simulated data with the known scintillation perturbations, we can also check the efficacy of the fit to the λ^2 terms as a technique for removing the RISS time delays. Averaging over the 5600 simulations, the procedure reduced the 1.4 GHz perturbations by an order of magnitude, from an rms amplitude ~ 1 μs at ~ 5 yr to $\lesssim 0.1$ μs. The 2.4 GHz perturbations were reduced to the same amplitude. Apparently, this simple prescription is quite sufficient, at the present level of timing accuracy, to measure and remove the significant contributions from interstellar propagation.

3. RESULTS AND CONCLUSIONS

In Figure 1 we show the power spectra of the observed and simulated DM corrections over the 5.2 yr of dual-frequency data for $m = 1$–8. The histograms, showing the dispersion of simulated results, are plotted along with the solid line connecting the mean values. The observed power spectrum is also shown. Note that irregularity in the data sampling prevents the

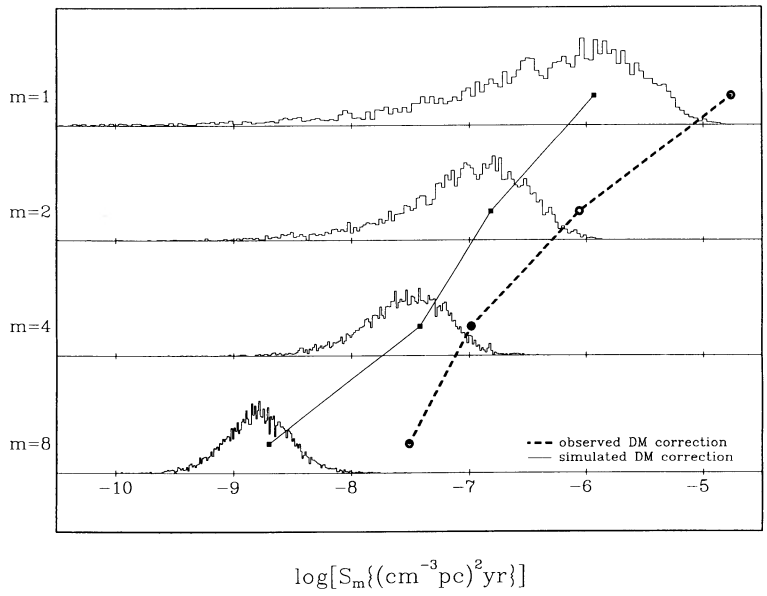

FIG. 1.—Power in the $\lambda^2(\delta\mathrm{DM})$ perturbations for a data subdivision m shown for the observed PSR 1937+21 data (*dashed lines*) and for average (*solid line*) and full distribution (*histograms*) of the simulated data.

m from corresponding exactly to frequencies, although the bandpass centers are reasonably close to m/T_{obs}.

Figure 1 shows that the observed λ^2 modulation in the $m = 1, 2, 4$ bins agrees in slope with that expected from a Kolmogorov spectrum of interstellar turbulence, although the amplitude is substantially lower than that of our nominal normalization. White-noise errors, from diffractive scintillation and finite precision measurements, will be independent in the two data sets—the resultant small DM fluctuations predicted from the observed ~ 0.3 μs rms errors fully explain the power observed for $m = 8$. Using the computed distributions for $m = 1, 2, 4$, we derive a maximum likelihood estimate for the renormalization factor of 5.5, with 68% ("1 σ") upper and lower limits of 15 and 2.5, respectively. The observed DM fluctuations accord well with the simulated Kolmogorov model histograms after these have been rescaled (shifted right) by this renormalization factor, although the best-fit power law appears somewhat steeper than $\beta = 11/3$.

In interpreting this renormalization, we note that in addition to the small factor uncertainty above due to diffractive scale definitions, there are appreciable uncertainties in the RISS amplitude due to the unknown scattering geometry. In particular, equation (2) shows that the screen normalization will scale as

$$\sim \left[\frac{D_p}{D_s(D_p - D_s)\Delta v_d} (v_{\rm rel}\tau)^2 \right]^{(\beta-2)/2} . \quad (5)$$

Note that scintillation observations (Cordes 1986) suggest a relative velocity for PSR 1937+21 somewhat higher than 50 km s^{-1}, although the true proper motion is quite small (Rawley et al. 1988). Also, if the scattering medium is on average less than halfway to the pulsar, the increase in RISS amplitude for the observed diffractive parameters can be quite considerable. Indeed, given that toward PSR 1937+21 we look out from near the edge of the local spiral arm across the gap of an interarm region (as evidenced by neutral hydrogen maps), we might expect that the bulk of the turbulent ionized electrons lie relatively close to Earth. Thus, it appears that the relatively uncertain normalization of the refractive variations might easily be increased to the level of the observed fluctuations.

In Figures 2 and 3, we display the measured polynomial coefficient power spectra and simulated power distributions for the arrival time residuals at 1.4 GHz and 2.4 GHz, respectively. We have rescaled the simulations by the maximum likelihood factor above. Several important conclusions can be derived from these results. First, scintillation effects are strong at the lower frequency and the observed powers are in reasonable agreement with simulated distributions for $m = 2, 4, 8$. However, for $m = 1$, the longest time scales accessible in the 7.1 yr data span, there appears to be appreciable excess noise. The probability of RISS-induced residuals having a power in this band as large or larger than that observed is only 0.009 with our maximum likelihood normalization, 0.12 with the upper limit. At 2.4 GHz (Fig. 3), the expected amplitude of the RISS timing residuals is appreciably smaller. For $m = 2, 4, 8$, the noise amplitude is in fact consistent with the value expected from the white-noise measurement errors seen at even higher noise frequencies. For $m = 1$, however, the data show a clear excess above the value predicted by the RISS simulations; the simulations show that such large powers have a probability $\sim 6 \times 10^{-4}$ of being RISS-induced for a $\beta = 11/3$ spectrum (maximum likelihood normalization), or $\sim 3 \times 10^{-3}$ ("1 σ" upper limit). Since at 2.4 GHz the red noise is significant only in the $m = 1$ ($\sim 1/5.2$ yr bin), we compare this power with that measured at 1408 MHz by interpolating between the 1.4 GHz $m = 1$ ($\sim 1/7.1$ yr) and $m = 2$ ($\sim 2/7.1$ yr) bins; the red-noise power is found to be roughly independent of the radiofrequency of observation.

However, in noise frequency, the spectrum of this additional power is substantially steeper than that of the RISS perturbations. Taking the $m = 1, 2$ points at 1.4 GHz and the $m = 1$ point of the 2.4 GHz data set and subtracting the power expected from scintillation alone, we see that the additional

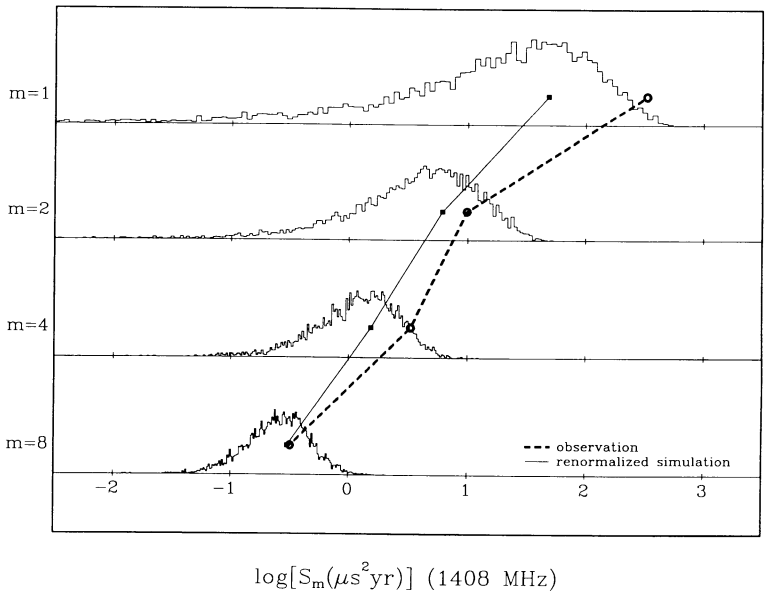

FIG. 2.—Power in the arrival time perturbations for a data subdivision m shown for the observed 1408 MHz PSR 1937+21 data (*dashed lines*). The distribution (*histograms*) and average (*solid line*) of the powers from the renormalized simulation are also shown.

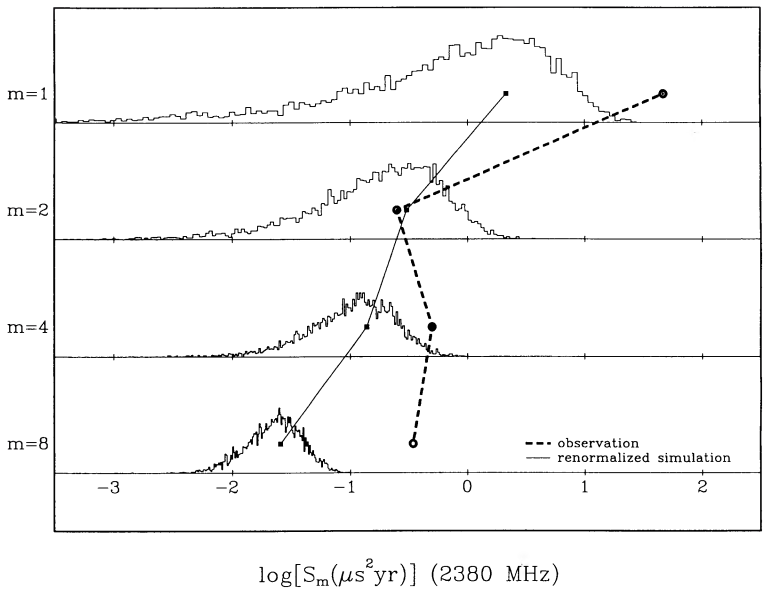

FIG. 3.—As in Fig. 2 for the 2380 MHz data; clear excess over the power expected from RISS (*histograms*) is observed (see text)

power has a nominal spectral index $\sim 6 \pm 1$ (formal errors). For comparison, the RISS spectra are predicted to have spectral indices ~ 3 (with the DM results giving a slightly higher value). Note that two candidates for this residual noise source, intrinsic instabilities in the pulsar spin and a stochastic background of cosmological gravitational waves, have expected indices of ~ 2–4 and ~ 5, respectively.

These simulations may be exploited to further improve our understanding of refractive scintillation in PSR 1937+21 timing. We can analyze runs for other β to improve constraints on the power spectrum of interstellar turbulence. We might also evaluate the importance of a population of discrete interstellar refraction events (Fiedler et al. 1987; Romani, Blandford, & Cordes 1987). Using the simulations as a test bed, we can investigate improved algorithms for removal of the δDM fluctuations and assess the need for further frequencies in the refraction removal (FC). Also we can numerically investigate predicted correlations between other observables, such as pulsar flux and angular size (Blandford & Narayan 1985; Romani et al. 1986) that may also prove of use in δDM corrections. The present result, the calibration of the propagation-induced residuals and the clear separation of an additional frequency-independent noise, has important implications for the future of the pulsar timing experiment. As this additional noise source apparently has a remarkably steep spectrum, it should continue to dominate over any propagation effects as the observation span lengthens. In addition, if this noise is peculiar to PSR 1937+21, continued timing of PSR 1855+09 (SRTR) will soon reach a sensitivity where a lower background limit could be set. However, should a similar steep-spectrum noise appear in PSR 1855+09, this would strongly motivate searches for angular correlations between the residuals. Such analyses are the critical discriminant between a gravitational wave background and other red-noise sources.

We wish to thank Jim Cordes, Roger Foster, and Joe Taylor for helpful discussions and the referee for a careful reading. This work was partially supported by NSF grants PHY86-20266 and AST88-17826, a grant from the Corning Glass Works (R. W. R.), and a Precision Measurement Grant from the National Institute of Science and Technology (D. R. S.).

REFERENCES

Armstrong, J. W. 1984, Nature, 307, 527
Armstrong, J. W., Cordes, J. M., & Rickett, B. J. 1981, Nature, 291, 562
Blandford, R., & Narayan, R. 1985, MNRAS, 213, 591
Blandford, R., Narayan, R., and Romani, R. W. 1984, J. Ap. Astr., 5, 369
Cordes, J. M. 1986, ApJ, 311, 183
Cordes, J. M., Pidwerbetsky, A., & Lovelace, R. V. 1986, ApJ, 310, 737
Cordes, J. M., Wolszczan, A., Dewey, R. J., & Blaskiewicz, M. 1990, ApJ, 349, 245
Fiedler, R., Dennison, B., Johnson, K., & Hewish, A. 1987, Nature 326, 675
Foster, R. S., & Cordes, J. M. 1990, ApJ, 364, 123 (FC)
Rawley, L. A., Taylor, J. H., & Davis, M. M. 1988, ApJ, 326, 947
Romani, R. W., Blandford, R., & Cordes, J. 1987, Nature, 328, 324
Romani, R. W., Narayan, R., & Blandford, R. D. 1986, MNRAS, 220, 19
Ryba, M., Taylor, J., & Stinebring, D. 1990, BAAS, 21, 1108
Stinebring, D., Ryba, M., Taylor, J., & Romani, R. W. 1990, Phys. Rev., 65, 285 (SRTR)
Taylor, J. H., & Weisberg, J. M. 1989, ApJ, 345, 434