Mon. Not. R. Astron. Soc. **288**, 533–537 (1997)

Search of a gravitational wave background in timing residuals of PSR 1937 + 21: minimal model and upper limits on Ω_{gr}

Gilles Zalamansky,[1] Christian Robert,[2] François Vernotte[1] and François Taris[3]

[1] *Besançon Observatory, BP 1615, 25010 Besançon, France*★
[2] *CREST, INSEE, timbre J340, 75675 Paris Cedex 14, France*
[3] *BNM-LPTF, Paris Observatory, 75014 Paris, France*

Accepted 1997 February 26. Received 1997 February 20; in original form 1996 August 5

ABSTRACT

Although many phenomena can account for pulsar timing residuals, previous upper limits on the amplitude Ω_{gr} of a gravitational wave background were all obtained with the same decomposition into a measurement noise with known variance and a gravitational noise parametrized by Ω_{gr}. The justification of this minimal model (MM) is its ability to provide, on average over an infinite set of measurements, the highest upper limits on Ω_{gr} compared with other models. Keeping with this model, and using free-access Arecibo data concerning PSR 1937 + 21, we derive a set of 157 estimators of Ω_{gr} and obtain the 95 per cent confidence upper limit of 1.7×10^{-7} using a Bayesian method. We also show evidence of a third term in the residuals (97 per cent confidence), which questions the adequacy of the MM.

Key words: cosmic strings – gravitation – methods: statistical – pulsars: individual: PSR 1937 + 21 – early Universe – large-scale structure of Universe.

1 INTRODUCTION

Cosmic strings which were possibly produced in the early Universe (Kibble 1976; Vilenkin & Shellard 1994) may have emitted a highly energetic gravitational wave background (GWB) in the range [10^{-10}, 10^{-6} Hz]. The relevant parameter to describe this homogeneous and isotropic background is its energy density per unit volume and its logarithmic frequency in units of closure density, $\Omega(f) = \rho_c^{-1}(\partial \ln \rho / \partial f)$.

Assuming a scale-invariant evolution (in horizon units) of a cosmic string network, Vachaspati & Vilenkin (1985) derived an analytic expression for $\Omega(f)$ which predicts

$$\Omega(f) \approx \Omega_{gr} \approx 10^{-7} \tag{1}$$

independently of f. This was derived soon after the discovery of the millisecond pulsar PSR 1937 + 21 (Backer, Kulkarni & Taylor 1983), the high stability of which enables us to set upper limits on Ω_{gr} when used as a clock in a Doppler effect detection experiment. It is of crucial interest in cosmology to set limits on Ω_{gr}, since cosmic strings could have introduced the primordial density contrast that eventually grew into the large-scale structures of the Universe (Hogan & Rees 1984; Vilenkin & Shellard 1994). Upper limits on Ω_{gr} are also interesting in time metrology (Taylor 1991). Although several authors (see Caldwell & Allen 1992 and Caldwell, Battye & Shellard 1996 for a detailed review) do not assume $\Omega(f)$ to be flat, we will keep with equation (1) in this paper, in order not to depart from the choices made by Stinebring, Kaspi, Taylor and Ryba (Stinebring et al. 1990; Kaspi, Taylor & Ryba 1994, KTR). The procedure we develop here can be easily extended to any parametrized $\Omega(f)$. We restricted ourselves to PSR1937 + 21 and we used the free Web access files 1937ao.92r (Kaspi et al. 1994).

2 PSR 1937 AND THE GRAVITATIONAL WAVE BACKGROUND

An ideal Doppler effect detector of gravitational waves consists of a perfectly stable clock $\mathscr{C}$ and a receiver $\mathscr{R}$ at a distance L, both freely falling with null reciprocal speed in a flat metric, clocks at $\mathscr{R}$ and $\mathscr{C}$ measuring proper times. Calling T_E the time of emission of a given pulse, T_o its arrival time at $\mathscr{R}$, $T_2 = T_E + L/c$ the expected arrival time under the assumption of no perturbation, and $R = T_o - T_2$ the timing residual, it can be shown (Mashoon 1979, 1982; Mashoon & Grishchuk 1980; Bertotti, Carr & Rees 1983) that if this residual is because of a GWB, its spectral density

★World Wide Web: http://www.obs/besancon.fr/www/tf/tf.html

$S_R(f)=\int\langle R(t)R(t+\tau)\rangle e^{i2\pi f\tau}\,d\tau$ can be linked with $\Omega(f)$: $S_R(f)=1.34\times10^4\Omega(f)f^{-5}B(fL/c)\ \mu s^2\,yr$, where B is a function which filters out frequencies lower than c/L. In the case of pulsars it can be set to 1 in the period range [≈ 1 d, 10 yr] corresponding to the data span.

To use PSR 1937 as a Doppler effect detector of GWB, we need a three step modelling as follows.

(i) A model for this clock, since it is not observable. In the case of PSR 1937 this model is (Kaspi et al. 1994)

$$N(T_2)=N_0+\nu T_2+\dot{\nu}\frac{T_2^2}{2} \tag{2}$$

pulse number n being emitted at $T_{2n}=N^{-1}(n)$. Parameters $\{N_0, \nu, \dot{\nu}\}$ are estimated under the assumption that the correct values should produce no residuals. We suppose that there is no ambiguity on n (Kaspi et al. 1994). Calling p_{cl} the set $\{N_0, \nu, \dot{\nu}\}$, we can express T_{2n} as a function of p_{cl}: $T_{2n}=T_2(n, p_{cl})$.

(ii) A model for the receiver. It is supposed to be at the Solar system barycentre (SSB), which requires a transformation of terrestrial arrival times $t_1, \ldots, t_N$ into barycentric ones $T_{o1}, \ldots, T_{oN}$. A rough model for T_o is (Blandford, Narayan & Romani 1984)

$$T_o=t-\frac{\boldsymbol{k}\cdot\boldsymbol{r}}{c}+\frac{\boldsymbol{k}\cdot\boldsymbol{r}-(\boldsymbol{r}^2)}{2Lc}-\frac{D}{f^2} \tag{3}$$

where $\boldsymbol{k}$ is the unit vector of the SSB–pulsar direction, and $\boldsymbol{r}$ is the position of the Earth with respect to the SSB. D/f^2 is the dispersion delay because of interstellar medium, D being adjusted by observations at two frequencies. The pulsar's equatorial coordinates α, δ, proper motions $\mu_\alpha=\dot{\alpha}\cos(\delta)$, $\mu_\delta=\dot{\delta}$, and parallax π are estimated under the assumption that the correct values should produce no residuals. We call p_a this set of astrometric parameters, $\{\alpha, \delta, \mu_\alpha, \mu_\delta, \pi\}$.

(iii) We need a model for the residuals $R_n=T_o(t_n, p_a)-T_2(p_{cl})$. Writing $R_n=\{1\}+\{2\}+\{3\}+\{4\}$, with $\{1\}=[T_o(t_n,p_a)-T_o(\bar{t}_n,p_a)]$, $\{2\}=\{T_o(\bar{t}_n,p_a)-T_o(\bar{t}_n,\bar{p}_a)]$, $\{3\}=[T_o(\bar{t}_n,\bar{p}_a)-T_2(\bar{p}_{cl})]$, $\{4\}=[T_2(\bar{p}_{cl})-T_2(p_{cl})]$, where bars mean 'true values', we see that R_n may come from errors in the measurement of the times of arrival, in the values given to the astrometric and clock parameters, and in the model itself $\{3\}$, at any of the three steps. GWB is a possible candidate, but so are departures of PSR 1937 + 21 from (2), drifts in Arecibo's clock, errors in ephemerids, etc.

2.1 TEMPO program and the least-squares fit of parameters

To estimate p_a and p_{cl}, the TEMPO program, which monitors pulsar timing at Arecibo, minimizes the sum (Kaspi et al. 1994)

$$\chi^2(p_a,p_{cl})=\sum_{n=1}^{N}\left(\frac{R_n}{\sigma_n}\right)^2=\sum_{n=1}^{N}\rho_n^2 \tag{4}$$

with $\sigma_1, \ldots, \sigma_N$ the uncertainties of each measurement $t_1, \ldots, t_N$.

In a linear approximation, this least-squares fit is equivalent to cancelling the projection of $\boldsymbol{\rho}$ over a subset H^σ of $\Re^N$ generated by the partial derivatives of (weighted) $\bar{T}_o$ and $\bar{T}_2$, with respect to p_a and p_{cl} (Blandford et al. 1984). The dimension of H^σ is 8 with models (2) and (3). [We did not include $\partial\bar{T}_2/\partial D$ into H^σ, since file 1937ao.92r only refers to frequency $f_1=2300$ MHz. Had it contained data corresponding to frequency $f_2=1400$ MHz, we would have included $\partial\bar{T}_2/\partial D=\boldsymbol{u}_1/f_1^2+\boldsymbol{u}_2/f_2^2$ into H^σ, with $u_{1i}=-1$ if $F(T_{oi})=f_1$, $u_{1i}=0$ if $f(T_{oi})=f_2$, and inversely for $\boldsymbol{u}_2$].

3 THE MINIMAL MODEL

We now write a general model for the redshift

$$\boldsymbol{\rho}=\boldsymbol{w}+\boldsymbol{g}+\boldsymbol{\rho}_{par}+\boldsymbol{\rho}' \tag{5}$$

$\boldsymbol{w}$ = weighted measurement noise, with covariance matrix $\mathscr{I}_N$ (identity matrix).

$\bar{\boldsymbol{g}}$ = weighted gravitational term, with covariance matrix $(\mathbf{V}_{gr})_{ij}=\langle g_i g_j\rangle$,

$$(\mathbf{V}_{gr})_{ij}=\frac{1.34\times10^{-3}\Omega_7}{\sigma_i\sigma_j}\int_{-\infty}^{\infty}f^{-5}e^{i2\pi f(t_i-t_j)}\,df=\Omega_7\bar{G}_{ij}. \tag{6}$$

$\boldsymbol{\rho}_{par}$ = errors in p_a and p_b. In the linear approximation $\boldsymbol{\rho}_{par}$ belongs to H^σ.

$\boldsymbol{\rho}'$ = all other terms.

Let $H^\sigma_\perp$ denote the subset of $\Re^N$ orthogonal to H^σ. We avoid all estimators of Ω_{gr} to be biased by ρ_{par} if we make them only depend on the projection $\boldsymbol{\rho}_\perp$ of $\boldsymbol{\rho}$ over $H^\sigma_\perp$.

The integral in (6) diverges at 0, but the inverse time travel $1/L$ is a low-frequency cut-off. The three largest eigenvalues of $\bar{\mathbf{G}}$ go to infinity when this cut-off goes to zero, whereas the others are only slightly affected by this. Since f_l is imprecisely known, wisdom advises us to include corresponding eigenvectors in H^σ. Fortunately they already belong to it, since they coincide with the partial derivatives of (weighted) $\bar{T}_o$ with respect to $\{N_o, \nu, \dot{\nu}\}$ (Blandford et al. 1984).

3.1 Construction of estimators

Kaspi et al. (1994) and before them Stinebring, Taylor, Ryba and Romani (Stinebring et al. 1990) (STR) used a minimal model (MM) for the residual, in which $\boldsymbol{\rho}'=\mathbf{0}$. This is supposed to yield the highest possible upper limits on Ω_7 allowed by the timing of PSR 1937. The probability distribution of $\boldsymbol{\rho}$ in the frame of the MM is (Borovkov 1987; Papoulis 1991)

$$p(\boldsymbol{\rho})=\frac{1}{(2\pi)^{N/2}\sqrt{\det\mathbf{V}}}\exp\left[-(\boldsymbol{\rho}-\boldsymbol{\rho}_{par})^t\mathbf{V}^{-1}(\boldsymbol{\rho}-\boldsymbol{\rho}_{par})\right] \tag{7}$$

with

$$\mathbf{V}=\mathscr{I}_N+\mathbf{V}_{gr}=\mathscr{I}_N+\Omega_7\bar{\mathbf{G}}.$$

A general treatment would consist of estimating the probability distribution of (p_a, p_{cl}, Ω_7) given $\boldsymbol{\rho}$, using a Bayesian method, and deducing the marginal distribution for Ω_7 by integrating over p_a and p_c. This implies an eight-dimensional integration, and the necessity of a precise choice of f_l. The projection of $\boldsymbol{\rho}$ over $H^{\sigma\perp}$ greatly simplifies the task, since Ω_7 is the only parameter in the distribution of $\boldsymbol{\rho}_\perp$.

In file 1937ao.92r, $N=165$. Using a Gram–Schmidt procedure we built an orthonormal basis $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N$ of $\Re$ with the first 157 generating $H^{\sigma}_{\perp}$ and the last eight H^{σ}. To derive this basis we used the DE200 ephemerides and the rough model (3), assuming that the differences between (3) and the true model used by TEMPO only introduces a negligible angle between our H^{σ} and the correct one.

Let $\eta_1, \ldots, \eta_N$ be the components of $\boldsymbol{\rho}$ over $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N$. In matricial formalism, with $\mathbf{A}^{\mathrm{t}}$ meaning the transpose of a matrix $\mathbf{A}$, this reads $\eta = \mathscr{H}^{\mathrm{t}}.\rho$, with $\mathscr{H}_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{h}_j$ [$\boldsymbol{e}_i$ = canonical basis of $\Re^N$, $\rho = (\rho_1, \ldots, \rho_N)^{\mathrm{t}}$].

Let $\mathbf{P}$ denote the rectangular $N_{\perp} \times N$ matrix, ($N_{\perp} = 157$, $P_{ij} = \delta_{ij}$) which performs a projection onto $H\sigma_{\perp}$. A straightforward calculation yields the covariance matrix $\langle \eta_{\perp}.\eta_{\perp}^{\mathrm{t}} \rangle$, ($\eta_{\perp} = \mathbf{P}.\eta$):

$$\langle \eta_{\perp}.\eta_{\perp}^{\mathrm{t}} \rangle = \mathbf{P}.\mathscr{H}^{\mathrm{t}}.(\mathscr{I}_N + \Omega_7 \bar{\mathbf{G}}).\mathscr{H}.\mathbf{P}^{\mathrm{t}} = \mathscr{I}_{157} + \Omega_7 \bar{\mathbf{W}}. \tag{8}$$

We now diagonalize $\bar{\mathbf{W}}$. We call $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{N_{\perp}}$ its eigenvectors, $\lambda_1, \ldots, \lambda_{N_{\perp}}$ its eigenvalues, and $\boldsymbol{q} = (q_1, \ldots, q_{N_{\perp}})$ the components of $\boldsymbol{\rho}$ over $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{N_{\perp}}$. We get the probability distribution of $\boldsymbol{q}$

$$p(\boldsymbol{q}) = \prod_{k=1}^{157} \frac{1}{\sqrt{2\pi(1+\lambda_k \Omega_7)}} \exp\left[-\frac{q_k^2}{2(1+\lambda_k \Omega_7)} \right] \tag{9}$$

which is the most simplified form of (7) we could hope to derive. We deduce $\boldsymbol{q}$ from $\boldsymbol{\rho}$ with the matricial product $q = \mathscr{W}^{\mathrm{t}}.\mathbf{P}.\mathscr{H}^{\mathrm{t}}.\rho$, $\mathscr{W}_{ij} = \boldsymbol{h}_i \cdot \boldsymbol{w}_j$.

In Fig. 1(a) we plot the eigenvalues $\lambda_1, \ldots, \lambda_{157}$ and in Fig. 1(b) the first eigenvector $\boldsymbol{w}_1$, which has some features of the vector $\boldsymbol{p}_3$ used by STR and KTR (Deeter & Boynton 1982; Deeter 1984) to derive their estimators S_1, S_2, S_4, S_8: $S_1 = (\boldsymbol{p}_3 \cdot \boldsymbol{R})^2$, $S_2, \ldots, S_8$ = averages of S_1 over one half-, …, one eighth of the data span. We plot in Fig. 2(a) the components $q_1, \ldots, q_{157}$, and in Fig. 2(b) the subset $q_{28}, \ldots, q_{157}$, whose variance is too large for those numbers to be $\mathscr{N}(0, 1)$ distributed, as we will now show.

3.2 Discussing the minimal model

Unless Ω_7 is large compared with 1, which is unlikely according to (Vachaspati & Vilenkin 1985; Bennett & Bouchet 1988; Caldwell & Allen 1992; Vilenkin & Shellard 1994; Caldwell et al. 1995), white noise dominates the distribution of all q_k for which λ_k is less than 1. The last 130 numbers ($q_{28}, \ldots, q_{157}$) are in such a situation, since $\lambda_{28}, \ldots, \lambda_{157}$ decrease from 0.015 to 10^{-6}. If we assume them to be Gaussian and standardized, then $S = q_{28}^2 + \ldots + q_{157}^2$ should be chi-square distributed with 130 degrees of freedom, which is close to $\mathscr{N}(130, 260)$. The number $(S-130)/\sqrt{260}$ should then be $\mathscr{N}(0, 1)$ distributed. The value we obtain, $s \approx 2.17$, makes this hypothesis quite unlikely: $|s|$ has probability 0.97 to be smaller under standard normal distribution. In other words the MM is not compatible with the data on PSR 1937 at the 97 per cent confidence level. Instead of writing $q_k = q_k^{(\mathrm{w})} + q_k^{(\mathrm{gr})}$, $\mathrm{Var}\,(q_k^{(\mathrm{w})}) = 1$, $\mathrm{Var}\,(q_k^{(\mathrm{gr})}) = \lambda_k \Omega_7$ where $q_k^{(\mathrm{w})}$ and $q_k^{(\mathrm{gr})}$ are the images of $\boldsymbol{w}$ and $\boldsymbol{g}$ by the projection $\mathscr{H}$ and the rotation $\mathscr{W}$, we should write $q_k = q_k^{(\mathrm{w})} + q_k^{(\mathrm{gr})} + q'_k$, with q'_k unknown.

On the average over an infinite set of measurements q_k^2 yields $E[q_k^2] = E[\alpha_k^2] + E[q_k'^2] + 2E[\varepsilon_k . q'_k]$, with $\varepsilon_k = q_k^{(\mathrm{w})} + q_k^{(\mathrm{gr})}$. Since q'_k is unlikely to depend on ε_k, then $E[q_k^2] = E[\varepsilon_k^2] + E[q_k'^2] = 1 + \lambda_k \Omega_7 + E[q_k'^2]$. If we use an MM in which $E[q_k^2] = E[\varepsilon_k^2] = 1 + \lambda_k \Omega_7$, we overestimate Ω_7 by a term $E[q_k'^2]/\lambda_k$. This is indeed what we expect from an MM. However, the probability to underestimate Ω_7 over one single measurement depends on the probability distribution of q'_k, which is obviously unknown. As the data span increases, λ_k will increase and the bias $E[q_k'^2]/\lambda_k$ will be reduced. At the present time improved estimates of Ω_7 definitely requires some information on q'_k.

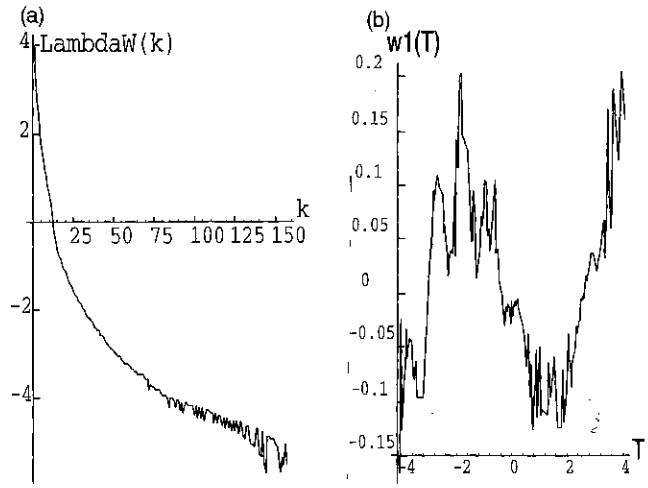

Figure 1. (a) Eigenvalues $\lambda_1, \ldots, \lambda_{157}$ of the matrix $\bar{\mathbf{W}}$. (b) The first eigenvector.

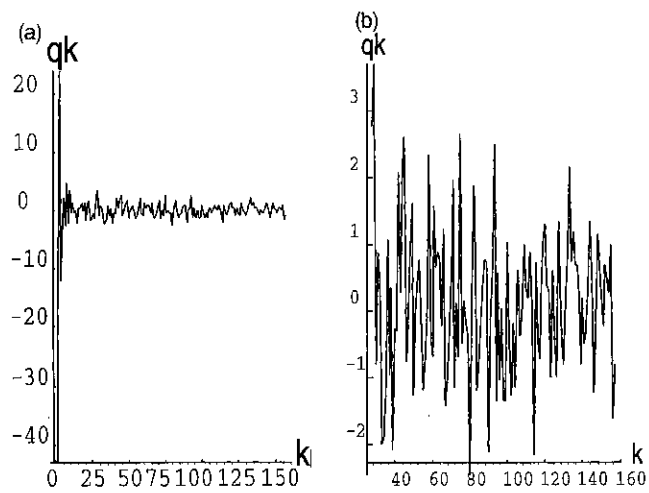

Figure 2. (a) Components of $\boldsymbol{\rho}$ over the basis $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{157}$ of $H^{\sigma}_{\perp}$. (b) The 'zoom' on the last 130 values.

4 BAYES THEOREM AND UPPER LIMITS ON Ω_7

Bayes theorem (Bayes 1763) tells us that the probability distribution of a measurement $\boldsymbol{x}$ given parameter(s) θ, and $\boldsymbol{x}$, is not enough to get a probability distribution for θ. One also needs an a priori distribution (the 'prior') $\pi(\theta)$, reflecting his or her degree of belief that θ should belong to given regions of the parameter space. The posterior distributon is then

$$p(\theta)=\frac{p(x|\theta)\pi(\theta)}{\perp p(x|\theta')\pi(\theta')\,\mathrm{d}\theta'}. \tag{10}$$

The everlasting debate is: which prior should we choose when we know nothing about θ? Two arguments favour 'Jeffreys prior'

$$\pi_{\mathrm{Jef}}(\theta)=\sqrt{|\mathscr{F}(\theta)|}, \tag{11}$$

$|\mathscr{F}(\theta)|$ is the determinant of Fisher's information matrix, $\mathscr{F}(\theta)_{ij}=\mathrm{Cov}\,(\partial \ln L/\partial\theta_i,\ \partial \ln L/\partial\theta_j)$, ($L$ being the likelihood). These arguments are as follows.

(1) The χ^2 distance between two distributions $p_1=p(x|\theta_1)$ and $p_2=p(x|\theta_2)$, with $\theta_2=\theta_1+\delta\theta$,

$$d_{\chi^2}=\int\left[\frac{p_2(x)-p_1(x)}{p_1(x)}\right]^2 p_1(x)\,\mathrm{d}x \tag{12}$$

is $\delta\theta^t.\mathscr{F}(\theta_1).\delta\theta+O(\|\delta\theta\|^2)$ (Borovkov 1987; Monfort 1996). d_{χ^2} can be considered as the average discriminating power of the measurement. In the parametrization $\phi(\theta)$ for which $\mathscr{F}(\phi)$ is constant [which is equivalent to (11) in a general parametrization (Robert 1994; Schervish 1995)], this discriminating power is the same for all ϕ. It is then sensible to choose a prior which respects this uniformity: $\pi(\phi)=1$ [which is also Laplace's prior (Laplace 1960)].

(2) When one performs a large number n of identical and independent measurements, the maximum likelihood estimator (MLE) distribution converges to a Gaussian one, with mean θ, and covariance matrix $[n\mathscr{F}(\theta)]^{-1}$ (Schervish 1995). With $\mathscr{F}(\phi)=1$ the MLE converges to the true value at uniform speed. See Bernardo (Bernardo 1979) and Robert (Robert 1994) for a more involved discussion.

It is straightforward to check that $\pi_{\mathrm{Jef}}(\theta)$ is uniform when θ is a translation parameter $[p(x|\theta)=\bar{p}(x-\theta)]$, that it is $1/\theta$ when θ is a scaling one $[p(x|\theta)=\bar{p}(x/\theta(/\theta]$, that it is invariant under reparametrization: $\pi_{\pi_{\mathrm{Jef}}}(\theta')=\pi_{\pi_{\mathrm{Jef}}}(\theta)|J(\theta)|$, $J(\theta)$ being the Jacobian determinant of the transformation, and that when the components of x are independent $\pi_{\mathrm{Jef}}(\theta)$ is unchanged under reordering. Any of those conditions, had it not been fulfilled, would have ruled out Jeffreys prior as a sensible choice.

In our case, all $Q_1,\ldots,Q_{157}$ with $Q_k=q_k^2$, are chi-square distributed:

$$p(Q|\Omega_7)=\prod_{k=1}^{157}\frac{1}{\sqrt{2\pi(1+\lambda_k\Omega_7)Q_k}}\,\mathrm{e}^{-Q_k/2(1+\lambda_k\Omega_7)}, \tag{13}$$

Fisher information is

$$\mathscr{F}(\Omega_7)=\sum_{k=1}^{157}\frac{\lambda_k^2}{(1+\lambda_k\Omega_7)^2},$$

and the prior we choose is

$$\pi(\Omega_7)=\sqrt{\mathscr{F}(\Omega_7)}\Upsilon(\Omega_7), \tag{14}$$

where Υ is the step function, equal to 1 over $[0,+\infty[$ and 0 over $]-\infty,0[$. With (14), (13) and (10) we obtained the probability density function and the corresponding cumulative density function for Ω_7. We plot them in Fig. 3(a) and (b). The upper limit at the 95 per cent confidence level we find for Ω_7 is 1.7×10^{-7}, which is slightly more stringent than Stinebring's value of 4×10^{-7} (Stinebring et al. 1990).

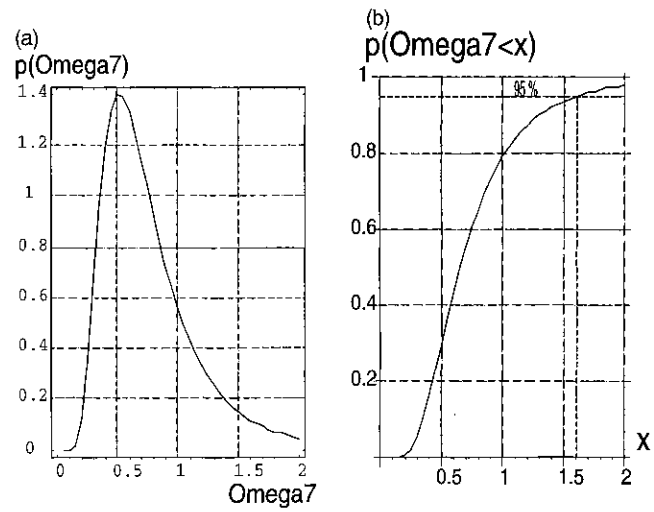

Figure 3. (a) Probability density function for $\Omega_7=10^7\Omega_{\mathrm{gr}}$. (b) Cumulative density function for Ω_7. We find that $0.95=R^{-1}(1.7)$, which is equivalent to concluding that 1.7×10^{-7} is the upper limit at the 95 per cent confidence level for Ω_{gr}.

4.1 PSR 1855 + 09

Using Stinebring's estimators (Stinebring et al. 1990) S_1, S_2, S_4, S_8, and their assumed probability distributions (Kaspi et al. 1994) (conditional to Ω_{gr}), the same Bayesian analysis (McHugh et al. 1996) yields the 95 per cent confidence upper limit 0.93×10^{-8} on Ω_{gr}. However, the construction of those estimators is not clearly detailed in papers Stinebring et al. (1990) and Kaspi et al. (1994), nor is the choice of two different clock models for both pulsars: least-squares of fit of TEMPO removes the cubic trend in timing residuals of PSR 1855 + 09 (Ryba & Taylor 1991), whereas it only removes the quadratic one in those of PSR 1937 + 21. This may explain why the values of S_2 and S_2 are so small compared with the assumed level of white noise (Kaspi et al. 1994). We thus restricted ourselves to PSR 1937 + 21.

4.2 Remark: tests of hypotheses

Consider only one estimator Q, the 'null' hypothesis $H(\omega):\Omega_7\geq\omega$ and the 'alternative', $\Omega_7=0$, where $[0,Q]$ is taken as rejection region for $H(\omega)$. The probability of rejecting $H(\omega)$ whereas it is true, is $\alpha(\omega)=p([0,Q]|\Omega_7\geq\omega)$ (type I error), and the probability of accepting it whereas it is false (type II error) is $\beta(\omega)=p([Q,\infty[|\Omega_7=0)$. Confusing $1-\alpha(\omega)$ with the probability for Ω_7 to be less than ω is misleading here, for the following three reasons. First, we would find a non-zero probability for Ω_7 to be less than 0 with this method. Secondly, if the $\sigma_1,\ldots,\sigma_N$ grow to infinity, Q and λ tend to 0, and so does $\alpha(\omega)$, for any ω, including very small values. In other words, the less reliable the measurements, the tighter the constraint on Ω_7, which is of course unacceptable! Thirdly, estimation and decision are two different problems: before taking a decision, one wants to estimate the probability to take the wrong one, α or β,

which is different from the probability for the parameter to be in a given region. This latter is related with $\pi(\theta)$, and this distribution is the starting point of Bayesian methods. Two authors (Thorsett & Dewey 1996) who recently used a Neyman Pearson test to set limits on Ω_7 obtained, we believe (McHugh et al. 1996), inadequate results.

5 CONCLUSION

An upper limit around 10^{-7} rules out the simplest models of structure formations based on cosmic strings (Turok 1986), yet not all of them (Vilenkin & Shellard 1994). We wanted to underline the model-dependence of this result, and the incompatibility of the generally assumed minimal model with the available data. Alternative models should be studied at this stage, a task we could not undertake in this paper.

REFERENCES

Backer D. C., Kulkarni S. R., Taylor J. H., 1983, Nat, 301, 314
Bayes T., 1763, Phil. Trans. Roy. Soc., 53, 370
Bennett D., Bouchet P., 1988, Phys. Rev. Lett., 60, 257
Bernardo J., 1979, J. Roy. Stat. Soc. B, 41, 113
Bertotti B., Carr B. J., Rees M. J., 1983, MNRAS, 203, 945
Blandford R., Narayan R., Romani R. W., 1984, JA&A, 5, 369
Borovkov A., 1987, Mathematical Statistics. Mir, Moscow
Caldwell R. R., Allen B., 1992, Phys. Rev. D, 45, 3447
Caldwell R. R., Battye R., Shellard E. P. S., 1996, proc. Topological Defects in Cosmology. Roma, in press
Deeter J. E., 1984, ApJ, 281, 482
Deeter J. E., Boynton P. E., 1982, ApJ, 261, 337
Hogan C. J., Rees M. J., 1984, Nat, 301, 109
Kaspi V. M., Taylor J. H., Ryba M. F., 1994, ApJ, 428, 713 (KTR)
Kibble W., 1976, J. Phys A, 9, 1387
Laplace P. S., 1960, Œuvres Complètes, Gauthier-Villars, Paris
Mashoon B., 1979, ApJ, 227, 1019
Mashoon B., 1982, MNRAS, 199, 659
Mashoon B., Grishchuk L. P., 1980, ApJ, 236, 990
McHugh M. P., Zalamansky G., Vernotte F., Lantz E., 1996, Phys. Rev. D, 54, 5993
Monfort A., 1996, Cours de Statistiques. Presses de l'Ecole Polytechnique, Paris
Papoulis A., 1991, Probability, Random Variables, and Stochastic Processes. McGraw-Hill, London
Robert C. P., 1994, The Bayesian Choice. Springer Verlag, New York
Ryba M. F., Taylor J. H., 1991, ApJ, 371, 739
Schervish M. J., 1995, Theory of Statistics. Springer Verlag, New York
Stinebring D. R., Ryba M. F., Taylor J. H., Romani R. W., 1990, Phys. Rev. Lett., 65, 285 (STR)
Taylor J. H., 1991, Proc. IEEE, 79, 1054
Thorsett S. E., Dewey R. J., 1996, Phys. Rev. D, 53, 3468
Turok N., 1986, Phys. Rev. Lett., 55, 1801
Vachaspati T., Vilenkin A., 1985, Phys. Rev. D, 31, 3052
Vilenkin A., Shellard E. P. S., 1994, Cosmic Strings and Other Topological Defects. Cambridge Univ. Press, Cambridge