*J. Astrophys. Astr.* (September 2017) 38:40
DOI 10.1007/s12036-017-9474-5



*Review*

# A New Look at Distances and Velocities of Neutron Stars

FRANK VERBUNT* and ERIC CATOR

Institute of Mathematics, Astrophysics and Particle Physics, Radboud University, P.O. Box 9010, 9500 AL Nijmegen, The Netherlands.
*Corresponding author. E-mail: F.Verbunt@astro.ru.nl

MS received 5 May 2017; accepted 4 August 2017; published online 7 September 2017

**Abstract.** We take a fresh look at the determination of distances and velocities of neutron stars. The conversion of a parallax measurement into a distance, or distance probability distribution, has led to a debate quite similar to the one involving Cepheids, centering on the question whether priors can be used when discussing a single system. With the example of PSR J0218+4232, we show that a prior is necessary to determine the probability distribution for the distance. The distance of this pulsar implies a gamma-ray luminosity larger than 10% of its spindown luminosity. For velocities, the debate is whether a single Maxwellian describes the distribution for young pulsars. By limiting our discussion to accurate (VLBI) measurements we argue that a description with two Maxwellians, with distribution parameters $\sigma_1 = 77$ and $\sigma_2 = 320$ km/s, is significantly better. Corrections for galactic rotation, to derive velocities with respect to the local standards of rest, are insignificant.

**Keywords.** Neutron stars—parallaxes—proper motions.

## 1. Introduction

This paper summarizes some of the results of a new look at pulsar distances (Igoshev *et al.* 2016) and pulsar velocities (Verbunt *et al.* 2017). We add some explanation and some illustrative computations.

The determination of distances to neutron stars is important because it forms the basis of the determination of their spatial density, and through this, of their birth rate. This in turn has consequences for our ideas about the progenitors of neutron stars, in particular, for the question of the lowest possible mass for a neutron star progenitor (e.g. Blaauw 1985; Hartman *et al.* 1997). Because of this importance, various indirect methods have been developed to establish distances, in addition to the direct geometric method of parallax measurement. In section 2, we compare the frequentist and Bayesian approaches to the determination of distance from a parallax measurement, to show that priors contribute significantly to the accuracy of the analysis.

In section 3, we take a brief look at a method for the distance determination that uses the dispersion measure and the luminosity function (cf. Verbiest *et al.* 2012). To derive a distance from the dispersion measure requires a model for the galactic electron-density distribution, and its accuracy depends critically on this model. It follows that the method should be used with care, as underestimation of errors may directly affect the conclusions drawn. In section 4, we compare the proper motions determined from timing with those determined from VLBI interferometry. In our description of the velocity distribution of young pulsars, we limit ourselves to pulsars for which distance and proper motion are derived from accurate VLBI measurements (section 5). We briefly discuss simple indications that the previously derived distribution, approximated by a Maxwellian with distribution parameter $\sigma \simeq 265$ km/s (Hobbs *et al.* 2005) is not acceptable. We then apply a full analysis to show that a description with the sum of two Maxwellians does better justice to the observation of a relatively large number of pulsars with low velocities (section 6).

## 2. Distance from parallax

Faucher-Giguère & Kaspi (2006), in their investigation of the birth velocity of pulsars, give an equation (their equation (2)) that converts the uncertainty of the parallax measurement into the uncertainty of the distance. This equation is in serious error,

as a result of confusion between the frequentist and Bayesian approaches to the treatment of measurement errors. (We explain this in more detail, in section 2.3). A similar error is made (their equation (3)) in the conversion of uncertainty in the dispersion measure to the uncertainty of the distance (as detailed in section 3). Unfortunately, these errors have been repeated in several later papers by Verbiest *et al.* (2010, 2012, 2014).

Incidentally, the confusion between the frequentist and Bayesian approaches is also in evidence in the study of Cepheid distances, in a slightly different form. Several authors, even in fairly recent papers, state that the parallax of a single object is not biased (e.g. Feast 2002, Francis 2014). This is all the more surprising as the correct treatment is well known, as explained in among others, Brown *et al.* (1997), Sandage & Saha (2002), and more recently by Bailer-Jones (2015).

Brown *et al.* (1997) also point out that the Lutz–Kelker effect (Lutz & Kelker 1973) must be applied with care. In its original form, this effect is computed with the assumption that the sources are distributed homogeneously throughout space, leading to an *a priori* probability of distance increasing with the square of the distance. For galactic sources, this assumption does not apply, and a distance prior must be constructed for each class of objects separately. In the case of pulsars, Verbiest *et al.* (2012) determined an appropriate prior for the distances.

### 2.1 *Frequentist and Bayesian treatment of measurement errors*

Let us, for simplicity, assume that the probability $g(p'|p)$ of measuring parallax $p'$ when the actual parallax is $p$ is given by a Gaussian

$$g(p'|p)\mathrm{d}p' = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(p'-p)^2}{2\sigma^2}\right) \mathrm{d}p', \qquad (1)$$

where $\sigma$ indicates the measurement error (Fig. 1). For an actual parallax $p$, this implies that in 68% of the cases

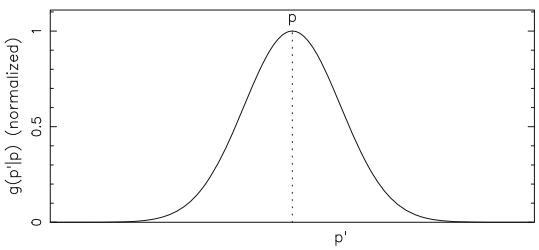


**Figure 1.** The probability of measuring $p'$ when the actual value is $p$ for the gaussian case with measurement error $\sigma$.

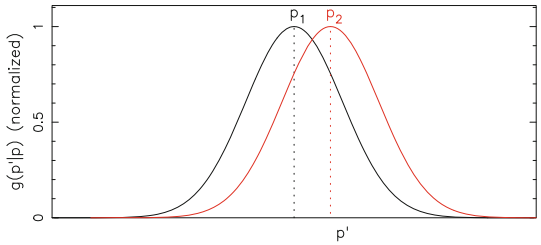


**Figure 2.** The probability of measuring $p'$ when the actual value is $p_1$ (black) or $p_2$ (red) for the gaussian case.

$|p'-p| < \sigma$, i.e. $p'-\sigma < p < p'+\sigma$. Now consider the measurements for two different actual parallaxes, $p_1$ and $p_2$ (Fig. 2). For each, we have

$$p'-\sigma < p_1 < p'+\sigma \qquad (68\%),$$
$$p'-\sigma < p_2 < p'+\sigma \qquad (68\%).$$

The intervals are the same even when $p_1$ and $p_2$ are different. More generally for any $p_i$,

$$p'-\sigma < p_i < p'+\sigma \qquad (68\%).$$

Thus we can state that once a value $p'$ has been measured with measurement error $\sigma$, *the probability is* 68% *for any actual value* $p_i$ *that the actual value lies in the interval from* $p'-\sigma$ *to* $p'+\sigma$. More generally, for each probability we can determine a corresponding interval for $p_i$. For example, there is a 90% probability that $1.45(p'-\sigma) < p_i < 1.45(p'+\sigma)$. Hence the name frequentist for this approach. *However, from the measurement alone we have no information on the probability distribution within this interval.*

To obtain this information, we must know how many actual objects there are with $p_1, p_2, \ldots, p_i$, i.e. we must know the distribution $f(p)$ of $p$. After all, a given measurement $p'$ may result from any of the many actual values $p$, according to equation (1). The joint probability $P(p, p')$ of actual value $p$ and measured value $p'$ is given by

$$P(p, p')\mathrm{d}p\mathrm{d}p' = f(p)\mathrm{d}p\, g(p'|p)\mathrm{d}p' \qquad (2)$$

and the probability $P(p|p')$ of an actual value $p$ in an interval $\Delta p$ for a measured value $p'$ is found from this by normalizing over all possibilities:

$$P(p|p')\Delta p = \frac{f(p)g(p'|p)\Delta p}{\int_p f(p)g(p'|p)\mathrm{d}p}, \qquad (3)$$

where the denominator acts as a normalization constant. In this Bayesian approach, $f(p)$ is the prior for $p$.

To apply this to distances, we rewrite equation (1) in terms of the distance $D = 1/p$:

$$g_{\rm D}\left(p'\Big|\frac{1}{D}\right)\mathrm{d}p' = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(p'-1/D)^2}{2\sigma^2}\right)\mathrm{d}p'. \tag{4}$$

Note that in this equation, $p = 1/D$ is fixed, and that the variable is $p'$. Hence, in converting equation (1) into equation (4) no $\mathrm{d}p/\mathrm{d}D$ term is warranted. For the *a priori* distance distribution $f_{\rm D}(D)$, with $f_{\rm D}(D)\mathrm{d}D = f(p)\mathrm{d}p$ (conservation of numbers), we obtain the probability of actual distance $D$ when parallax $p'$ is measured as

$$P(D|p')\Delta D = \frac{f_{\rm D}(D)g_{\rm D}\left(p'|\frac{1}{D}\right)\Delta D}{\int_p f_{\rm D}(D)g_{\rm D}\left(p'|\frac{1}{D}\right)\mathrm{d}D}. \tag{5}$$

### 2.2 *The distance of PSR J0218+4232*

Igoshev *et al.* (2016) illustrated this last equation with the case of the millisecond pulsar PSR J0218+4232 (see Fig. 3). The distance prior is taken from Verbiest *et al.* (2012), and reflects the fact that we are looking from a location $R_{\rm o} = 8.5\,$kpc from the galactic center at a distribution around this center in the radial direction, and around the galactic plane in the vertical ($z$) direction. This leads to (in notation slightly altered from that in Verbiest *et al.* 2012):

$$f_{\rm D}(D)\mathrm{d}D = D^2R^{1.9}\exp\left(-\frac{D\sin b}{0.5\,\mathrm{kpc}} - \frac{R}{1.7\,\mathrm{kpc}}\right), \tag{6}$$

where $R$ is the distance of the pulsar to the galactic center, projected on the galactic plane:

$$R = \sqrt{{R_{\rm o}}^2 + (D\cos b)^2 - 2D\cos b\, R_{\rm o}\cos l}. \tag{7}$$

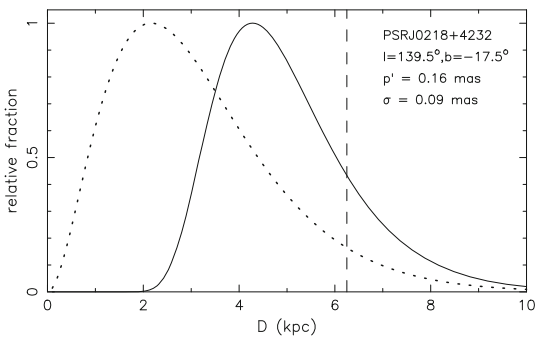


**Figure 3.** The *a priori* distribution $f_{\rm D}(D)$ of millisecond pulsars in the direction of PSR J0218+4232 (dotted line, equations (6), (7)), and the measured parallax $p' = 0.16 \pm 0.09\,$mas, lead to the distance probability distribution given by the solid line, according to equation (5). The vertical dashed line indicates the nominal distance $D' = 1/p'$.

This prior is shown in Fig. 3 as a dotted line, for the direction of PSR J0218+4232. Equation (4) shows that a measured parallax $p'$ can result from a range of distances; the probability that a measured $p'$ is due to an actual distance $D$ scales with the product of equation (4) with the number $f_{\rm D}(D)$ of objects at that distance $D$. After normalization, this leads to the probability density function expressed in equation (5), and is shown for PSR J0218+4232 in Fig. 3.

The factor $g_{\rm D}$ in equation (5) leads to a shift of the most probable value of distance $D$ from the peak of $f_{\rm D}(D)$ to values closer to the distance $D' = 1/p'$. Conversely, the factor of the prior $f_{\rm D}(D)$ leads to a shift of the most probable value of $D$ from the nominal distance $D'$ towards the peak of the prior distribution.

In the basic form of the Lutz–Kelker effect, for a homogenous distribution $f_{\rm D}(D) \propto D^2$, the most probable actual distance is always larger than the nominal distance $D' = 1/p'$. Figure 3 illustrates the fact that the Lutz–Kelker effect in a more general form, i.e. allowing other forms of $f_{\rm D}(D)$, may cause the most probable distance to be *lower* than the nominal one.

### 2.3 *Confusing frequentist and Bayesian approaches*

For a flat prior, $f_{\rm D}(D) =$ constant, equations (4) and (5) simplify to

$$\begin{aligned} P(D|p')\Delta D &\propto g_{\rm D}\left(p'\Big|\frac{1}{D}\right)\Delta D \\ &= \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(p'-1/D)^2}{2\sigma^2}\right)\Delta D. \end{aligned} \tag{8}$$

This equation is very similar to equation (4), but there is a crucial difference: the probability of equation (4) is normalized by integrating over $p'$, the probability of equation (8) is normalized by integrating over $D$. Misreading equation (8) as valid for an interval $\Delta p'$ leads one to write $\Delta p' = (1/D^2)\Delta D$, and thereby add a Factor $1/D^2$ to equation (4). It appears that this is what Faucher-Giguère & Kaspi have done.

In fact, as may be seen from equation (5), this corresponds to assuming a prior $f_{\rm D}(D) \propto 1/D^2$.

## 3. Distance from dispersion measure or luminosity

In principle, the dispersion measure $DM$, giving the integrated number of electrons between Earth and the pulsar, can be combined with a model for the electron distribution in the Milky Way, to determine the pulsar distance. It is well known that this method gives

rather uncertain, and occasionally clearly wrong results (e.g. for B1929+10, see Table 5 in Brisken *et al.* 2002). Brisken *et al.* (2002) followed by Faucher-Giguère & Kaspi (2006) and by Verbiest *et al.* (2012), tried to circumvent this problem by 'assigning the $DM$ a gaussian probability distribution function centered on the measured value $DM_o$ with a 40% variance':

$$f_{\rm DM}(DM) \propto \exp\left[-0.5\left(\frac{DM - DM_o}{0.4DM_o}\right)^2\right]. \quad (9)$$

This provides a rough guess of the uncertainty of a distance derived from $DM$ and a model electron distribution.

In principle, even large measurement uncertainties lead to the correct result, if the measurements are properly weighted. Equation (9) simplifies the complexity of the galactic electron distribution too much to provide such proper weighting. Note, for example, that the probability for $DM = 0$ (hence $D = 0$) is non-zero, and indeed the same for all values of $DM_o$, no matter how large. Equation (9) suggests that the error in a distance derived from the dispersion measure is gaussian, where in fact the error is systematic: an error in the electron density model leads to a systematic shift in the derived distance.

Faucher-Giguère & Kaspi (2006), followed by Verbiest *et al.* (2012), compound the error by adding a multiplication factor $\mathrm{d}DM/\mathrm{d}D$ in equation (9), making an error analogous to the one for distances discussed in section 2.3. This factor has the clearly unphysical effect of concentrating the distance probability in areas of enhanced electron density, since $\mathrm{d}DM/\mathrm{d}D \propto n_e$.

Verbiest *et al.* (2012) also used the luminosity function to constrain the distance: the luminosity function peaks at low luminosities, hence a pulsar with a given flux is more likely a nearby low-luminosity one than a faraway bright pulsar. In converting a likelihood of luminosity $L$ into a likelihood of distance, Verbiest *et al.* (2012) erroneously introduce a $\mathrm{d}\log L/\mathrm{d}D$ factor. Igoshev *et al.* (2016) corrected this and showed that a wide variety of gamma-ray luminosity functions led to an isotropic gamma-ray luminosity in excess of 10% of the spindown luminosity for PSR J0218+4232.

Because of the steepness of the luminosity function, straightforward application of the resulting bias pushes the distance probability to the lowest distances allowed by other indicators. Our knowledge of the luminosity function of pulsars depends on our knowledge of distances, and thus, in principle, the luminosity function and distance distribution of pulsars should be determined together.

## 4. Velocity from timing and dispersion measure

The annual variation in the difference between heliocentric and geocentric pulse arrival times depends on the celestial position of the pulsar. This dependence may be used to determine the position of the source, and over time its parallax and proper motion, from pulse timing. Hobbs *et al.* (2005) listed a large number of proper motions for pulsars determined with this method. By comparing these proper motions and their uncertainties with the measurements for the same pulsars obtained with VLBI (Chatterjee *et al.* 2009; Brisken *et al.* 2002; Kirsten *et al.* 2015), we see that the measurement errors given for young (i.e. not recycled) pulsars are of the order a hundred times larger for timing measurements than for VLBI. Because of these large uncertainties no timing parallaxes have been determined for young pulsars.

Hobbs *et al.* (2005) therefore used distances estimated from dispersion measure to convert the proper motions into velocities. Their use of a non-parametric `clean` algorithm to determine the intrinsic velocity distribution, has the advantage of obviating the need to prescribe a parametrized form of this distribution. However, Hobbs *et al.* (2005) noted that the result is well described by a Maxwellian with distribution parameter $\sigma = 265$ km/s, and argued that the low values of velocity perpendicular to the line-of-sight observed for some pulsars are the result of projection effects.

One of the authors, Frank Verbunt always found it hard to accept this, for the following reason. An isotropic Maxwellian may be considered as composed of three gaussians, in three mutually perpendicular directions. If we choose the line of sight as one direction, the two remaining directions are in the celestial plane, and the two gaussians lying in this plane may be combined to give the distribution of $v_\perp$. The fraction of velocities in this distribution below any $v_c$ may be written (for derivation, see equation (A.4) in Appendix)

$$f(v_\perp < v_c) = 1 - \mathrm{e}^{-{v_c}^2/(2\sigma^2)}. \quad (10)$$

Table 5 of Brisken *et al.* (2002) listed the nine accurate velocities $v_\perp$ known at the time, and out of these two have $v_\perp < 40$ km/s. For $\sigma = 265$ km/s and $v_c = 40$ km/s, the probability for one trial that $v_\perp < v_c$ follows from equation (10) is about 1%. The probability of finding 2 in 9 trials is 0.4%. This suggests that the fraction of low velocities is underestimated by the analysis of Hobbs *et al.* (2005). Remarkably, this original argument for the velocity study of Verbunt *et al.* (2017) was rather weakened when the accurate proper motion data for 28 pulsars were collected. Not a single

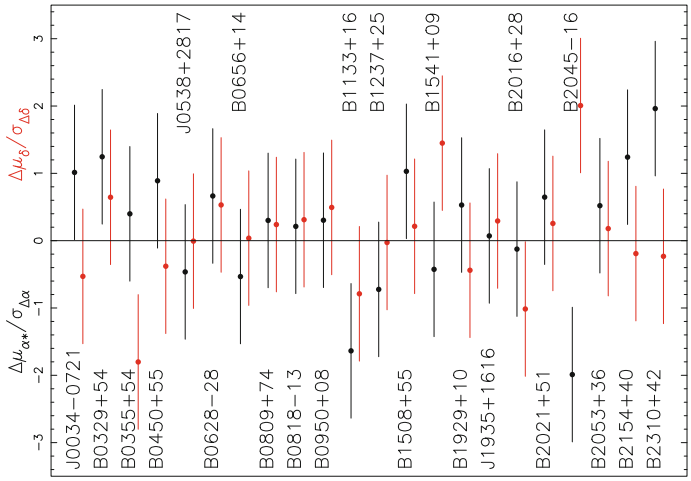


**Figure 4.** The difference $\Delta\mu_{\alpha*} = \mu_{\alpha*,\mathrm{VLBI}} - \mu_{\alpha*,\mathrm{tim}}$ between proper motions in the direction of right ascension $\mu_{\alpha*,\mathrm{VLBI}}$ and $\mu_{\alpha*,\mathrm{tim}}$ measured with VLBI and with timing, respectively, in units of the error in the difference $\sigma_{\Delta\alpha}$ (black); and analogous for the difference $\Delta\mu_\delta$ of the proper motions in the direction of declination (red).

new one with $v_\perp < 40$ km/s was added! The probability of finding 2 in 28 trials is about 4%.

As we will see below, a single Maxwellian *does* underestimate the number of low velocity pulsars, albeit at less low velocities than suggested by the two velocities below 40 km/s. Such an underestimation may arise if Hobbs *et al.* (2005) underestimated the velocity errors.

Figure 4 compares the proper motions determined from timing with those determined from VLBI, for pulsars with accurate VLBI measurements, by plotting the difference between the proper motions in units of the error in the difference, for the directions of right ascension and of declination separately. The figure shows that the errors for the timing proper motions, although large, are reliable, in the sense that they are distributed around the correct (VLBI) values as expected. Thus, in velocities determined with proper motions from timing and distances from dispersion measure, the problem for a reliable statistical analysis lies in the distances.

## 5. Velocity from VLBI measurements

Given the large errors in the velocities derived with distances from the dispersion measure and proper motions from timing, it appears appropriate to make a first effort at determining the velocity distribution on the basis of the smaller sample with VLBI parallaxes and proper motions. With these much smaller errors, exact understanding of the error distribution is less critical. We collect from the literature 28 young (in the sense of not recycled) pulsars for which these data are available. We indicate the measured values and the nominal values derived from them with a prime: parallax $p'$ and proper motions $\mu'_{\alpha*}, \mu'_\delta$; and nominal distance $D' = 1/p'$ and velocity perpendicular to the line-of-sight $v'_\perp = \sqrt{{\mu'_{\alpha*}}^2 + {\mu'_\delta}^2}/p'$.

In Fig. 5, we show the cumulative distribution of $v'_\perp$, together with the cumulative distribution according to equation (10) for $\sigma = 265$ km/s. The Kolmogorov–Smirnov test gives a probability of 0.0024 that the observed distribution is drawn from this distribution. It shows that the Maxwellian predicts too few pulsars with low velocities, up to several hundred km/s. Some caution is required in the interpretation of this result, because the observed distribution shown in Fig. 5 and used in the Kolmogorov–Smirnov test, ignores measurement errors.

In Fig. 5, we also show the absolute values of the nominal velocities $v'_\alpha = \mu'_{\alpha*}/p'$ and $v'_\delta = \mu'_\delta/p'$, together with their nominal errors. The median of $v_\perp$ is found by equating the cumulative distribution of equation (10) to 0.5:

$$1 - \mathrm{e}^{-{v_{\perp,\mathrm{m}}}^2/(2\sigma^2)} = 0.5 \Rightarrow v_{\perp,\mathrm{m}} = \sigma\sqrt{2\ln 2}. \qquad (11)$$

This median is also shown in Fig. 5. It is seen that the errors in the lower velocities are small, indicating that our conclusion from the Kolmogorov–Smirnov test on $v'_\perp$ is reliable. Also, only 7 of 28 pulsars have $v_\perp$ higher than the median velocity predicted by a Maxwellian with $\sigma = 265$ km/s.

Figure 5 strengthens our earlier suspicion that a single Maxwellian underpredicts the number of

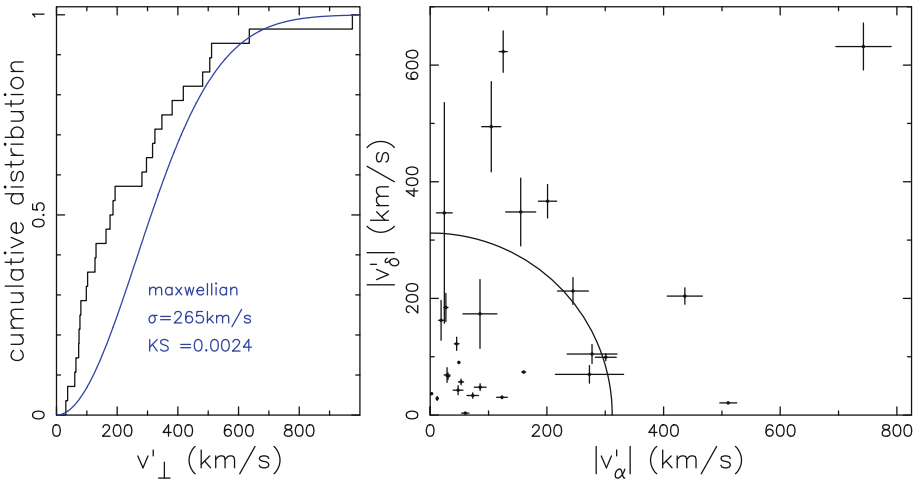


**Figure 5.** *Left*: The observed cumulative distribution of $v'_\perp$ derived from VLBI measurements for 28 pulsars compared with the cumulative distribution of $v_\perp$ (equation (10)) for a Maxwellian with $\sigma = 265$ km/s. *Right*: The nominal velocities $|v'_\alpha| = |\mu'_{\alpha*}|/p'$ and $|v'_\delta| = |\mu_\delta|/p'$. The circle gives the median velocity $v_{\perp,\mathrm{m}}$ (equation (11)) for $\sigma = 265$ km/s.

low-velocity pulsars. For a definite conclusion, however, we must perform an analysis which takes account of the measurement errors properly.

## 6. The interplay of distance, proper motion and velocity distribution

As a first prior for the intrinsic velocity distribution we consider a single isotropic Maxwellian. Each pulsar velocity is a draw from this Maxwellian, i.e. a draw from each of three gaussians in mutually perpendicular directions. For each pulsar, we choose the three directions along the line-of-sight and along the directions of increasing right ascension $\alpha$ and declination $\delta$, and thus for the direction along $\alpha$, we have the prior

$$f(v_\alpha,\sigma)\mathrm{d}v_\alpha = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-{v_\alpha}^2/2\sigma^2}\mathrm{d}v_\alpha \tag{12}$$

and analogously for $v_\delta$ and $v_r$. The joint probability of measured values for parallax and proper motions $p'$, $\mu'_{\alpha*}$ and $\mu'_\delta$, and actual distance and velocities $D$, $v_\alpha$, $v_\delta$ and $v_r$ follows as

$$\begin{aligned} P_{\mathrm{maxw}} &\equiv P_{\mathrm{maxw}}(p',\mu'_{\alpha*},\mu'_\delta,D,v_\alpha,v_\delta,v_r) \\ &= f_{\mathrm{D}} f(v_\alpha,\sigma) f(v_\delta,\sigma) f(v_r,\sigma) g_{\mathrm{D}} g_\alpha g_\delta, \end{aligned} \tag{13}$$

where $f_{\mathrm{D}}$ is given by equations (6), (7) and $g_{\mathrm{D}}$ by equation (4); $f(v_\alpha,\sigma)$ by equation (12), and $f(v_\delta,\sigma)$ and $f(v_r,\sigma)$ analogously; and $g_\alpha$ and $g_\delta$ by

$$g_\alpha = \frac{1}{\sigma_\alpha\sqrt{2\pi}}\exp\left[-\frac{(\mu_{\alpha*,\mathrm{G}}(D)+v_\alpha/D-\mu'_{\alpha*})^2}{2{\sigma_\alpha}^2}\right], \tag{14}$$

$$g_\delta = \frac{1}{\sigma_\delta\sqrt{2\pi}}\exp\left[-\frac{(\mu_{\delta,\mathrm{G}}(D)+v_\delta/D-\mu'_{\delta})^2}{2{\sigma_\delta}^2}\right], \tag{15}$$

where $\sigma_\alpha$ and $\sigma_\delta$ are the measurement errors in $\mu_{\alpha*}$ and $\mu_\delta$, respectively, and $\mu_{\alpha*,\mathrm{G}}(D)$ and $\mu_{\delta,\mathrm{G}}(D)$ the corrections due to galactic rotation, between the local standards of rest at the position of the Sun and the pulsar. These corrections are necessary, because we are interested in the peculiar velocity of the pulsar, not including the apparent velocity due to galactic rotation. Because most pulsars with an accurate parallax are nearby, these corrections generally are small.

To obtain the value of the scale parameter $\sigma$ which gives the most likely correspondence with the measurements, we must consider the contributions to the likelihood of all distances and velocities, i.e. integrate equation (13) over $D$, $v_\alpha$, $v_\delta$ and $v_r$. The integral over $v_r$ is 1; the integrals over $v_\alpha$ and $v_\delta$ are more involved, but can be done analytically. The resulting likelihood is (Verbunt *et al.* 2017):

$$\begin{aligned} L_{\mathrm{maxw}}(\sigma) &= \int_0^{D_{\max}}\int_0^\infty\int_0^\infty\int_0^\infty P_{\mathrm{maxw}}\,dD\mathrm{d}v_\alpha\mathrm{d}v_\delta\mathrm{d}v_r \\ &= \mathcal{C}\int_0^{D_{\max}} f_{\mathrm{D}} g_{\mathrm{D}} \mathcal{I}_\alpha \mathcal{I}_\delta \mathrm{d}D, \end{aligned} \tag{16}$$

where $\mathcal{C}$ is a constant, $D_{\max}$ the maximum distance (we use $D_{\max} = 10$ kpc; beyond this distance the factor $g_{\mathrm{D}}$ according to equation (4) ensures that the integrand is

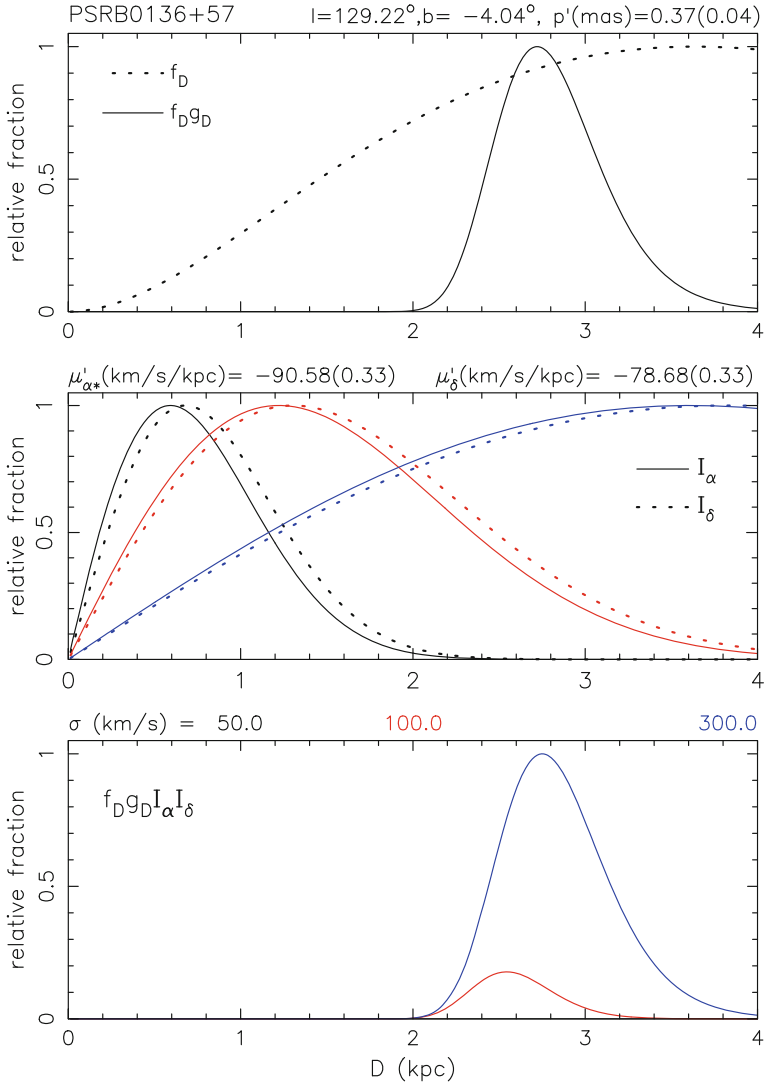


**Figure 6.** The integrand of equation (16). The distance implied by the parallax and galactic pulsar distribution in the direction of PSR B0136+57 (*top frame*), combined with the proper motion ($\mathcal{I}_\alpha$, $\mathcal{I}_\delta$, equation (17)). The *middle frame* implies a large velocity and favours a Maxwellian with high scale parameter $\sigma$. The curves in the top and middle frames are normalized to maximum value 1. The *bottom frame* shows the integrand of equation (16) is normalized such that the area under the curve is proportional to the likelihood $L_{\rm maxw}(\sigma)$ (equation (16)).

effectively zero for the pulsars in our sample), and we define

$$\mathcal{I}_\alpha \equiv \left(1+\frac{\sigma^2}{D^2{\sigma_\alpha}^2}\right)^{-1/2} \exp\left[-\frac{1}{2}\frac{(D\,\mu_{\alpha*,\rm G}-D\,\mu'_{\alpha*})^2}{\sigma^2+D^2{\sigma_\alpha}^2}\right],$$
$$\mathcal{I}_\delta \equiv \left(1+\frac{\sigma^2}{D^2{\sigma_\delta}^2}\right)^{-1/2} \exp\left[-\frac{1}{2}\frac{(D\,\mu_{\delta,\rm G}-D\,\mu'_{\delta})^2}{\sigma^2+D^2{\sigma_\delta}^2}\right]. \quad (17)$$

The effect of the separate contributors to the integrand of equation (16) is shown in Fig. 6, for the case of PSR B0136+57. The observational data $p'$, $\mu'_{\alpha*}$ and $\mu'_\delta$ for this pulsar are taken from Chatterjee *et al.* (2009). We convert the proper motion with

$$\mu({\rm km/s/kpc}) = 4.74\mu({\rm mas/yr}). \quad (18)$$

The accurate parallax and proper motion imply a velocity of several hundred km/s: the nominal projected velocity is $v'_\perp = 324\,$km/s. When we compare the probability of such a velocity for three different Maxwellians, with $\sigma = 50$, 100 and 300 km/s respectively, the probability of the one with $\sigma = 300\,$km/s is highest.The probability of the Maxwellian with $\sigma = 100\,$km/s is significantly lower, and the Maxwellian with $\sigma = 50\,$km/s is virtually excluded (its integrand invisible in Fig. 6).

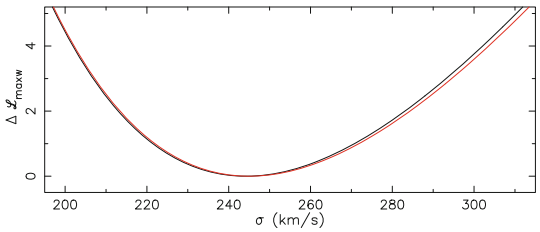


**Figure 7.** $\mathcal{L}_{\rm maxw}$ according to equation (19) as a function of the Maxwellian scale parameter $\sigma$. The black curve shows the result of the full calculation. The red curve, almost indistinguishable, shows the result when corrections $\mu_{\alpha*,\rm G}$ and $\mu_{\delta,\rm G}$ for galactic rotation are omitted.

### 6.1 *Description with a single Maxwellian*

To determine the best value of $\sigma$ for the complete set of 28 pulsars, Verbunt *et al.* (2017) first computed $L_{\rm maxw}(\sigma)$ according to equation (16) for each of them, integrating numerically over $D$. From these likelihoods, the deviance is computed as

$$\mathcal{L}_{\rm maxw}(\sigma) = -2\sum_{i=1}^{N} \ln L_{{\rm maxw},i}(\sigma), \tag{19}$$

where index $i$ labels the pulsar. With this definition of the deviance, the best value $\sigma_{\rm opt}$ is the one that minimizes $\mathcal{L}_{\rm maxw}$ (and thus maximizes the product of the likelihoods), and the differences

$$\Delta\mathcal{L}_{\rm maxw} \equiv \mathcal{L}_{\rm maxw}(\sigma) - \mathcal{L}_{\rm maxw}(\sigma_{\rm opt}) \tag{20}$$

approximate a $\chi^2$ distribution. $\Delta\mathcal{L}_{\rm maxw}$ is shown as a function of $\sigma$ in Fig. 7. The minimum of $\mathcal{L}_{\rm maxw}$ occurs at $\sigma_{\rm opt} \simeq 245$ km/s.

To see the effect of the corrections for galactic rotation to the observed proper motion, we also perform a calculation in which these corrections are omitted, i.e. in which $\mu_{\alpha*,\rm G}$ and $\mu_{\delta,\rm G}$ in equations (16) and (17) are put to zero. The result is the same within the uncertainty.

### 6.2 *Description with two Maxwellians*

As argued in section 5, a single Maxwellian is not a good description of the observed velocity distribution. To illustrate this,we show in Fig. 8 that the data for PSR B2016+28 (taken from Brisken *et al.* 2002) imply a low projected velocity: $v'_\perp = 31$ km/s. From the three Maxwellians considered, this velocity clearly favours the one with $\sigma = 50$ km/s.

As a second approach to the determination of the intrinsic velocity distribution of young pulsars, we therefore describe it with the sum of two Maxwellians:

$$f_v(\vec{\sigma}) = \sqrt{\frac{2}{\pi}} v^2 \left[ \frac{w}{\sigma_1^3} \exp\left(-\frac{1}{2}\frac{v^2}{\sigma_1^2}\right) + \frac{(1-w)}{\sigma_2^3} \exp\left(-\frac{1}{2}\frac{v^2}{\sigma_2^2}\right)\right] \tag{21}$$

with the parameter vector $\vec{\sigma} = w, \sigma_1, \sigma_2$. In analogy with equations (16), (19) and (20), we now have

$$L_{\rm 2maxw}(\vec{\sigma}) = wL_{\rm maxw}(\sigma_1) + (1-w)L_{\rm maxw}(\sigma_2), \tag{22}$$

$$\mathcal{L}_{\rm 2maxw}(\vec{\sigma}) = -2\sum_{i=1}^{N} \ln L_{{\rm 2maxw},i}(\vec{\sigma}), \tag{23}$$

$$\Delta\mathcal{L}_{\rm 2maxw}(\vec{\sigma}) \equiv \mathcal{L}_{\rm 2maxw}(\vec{\sigma}) - \mathcal{L}_{\rm 2maxw}(\vec{\sigma}_{\rm opt}). \tag{24}$$

Verbunt *et al.* (2017) computed $L_{\rm maxw}(\sigma)$ on a grid of $\sigma$ values with a spacing of 1 km/s, and use the `amoeba` routine from Press *et al.* (1986) to determine the values of $\vec{\sigma}_{\rm opt}$ that minimize $\mathcal{L}_{\rm 2maxw}$. They found that the best description of the velocity distribution is the combination of 42% of the pulsars in a Maxwellian with $\sigma_1 = 77$ km/s with a 58% in a Maxwellian with $\sigma_2 = 320$ km/s. Comparing the best solution for two Maxwellians with that for one Maxwellian, Verbunt *et al.* (2017) found $\mathcal{L}_{\rm 2maxw}(\vec{\sigma}_{\rm opt}) - \mathcal{L}_{\rm maxw}(\sigma_{\rm opt}) = -14$. For two added parameters, this difference indicates that the solution with two Maxwellians is significantly better.

The choice of $\mathcal{L}_{\rm 2maxw}$ according to equation (22) implies that the distribution of $\Delta\mathcal{L}_{\rm 2maxw}(\vec{\sigma})$ approximates a $\chi^2$ distribution. Thus we find the 68% and 95% probability contours in the $\sigma_1$ - $\sigma_2$ plane as delineated by $\Delta\mathcal{L}_{\rm 2maxw}(\vec{\sigma}) = 1$ and $\Delta\mathcal{L}_{\rm 2maxw}(\vec{\sigma}) = 4$, respectively. This is shown in Fig. 9.

To gauge the effect of the corrections for galactic rotation, we show in the same figure the results for a computation in which these corrections were set to zero. This leads to a marginal shift to a lower value (71 km/s) for $\sigma_1$. The value of $\sigma_2$ is not affected.

## 7. Conclusions

The distance derived from a parallax measurement of a single pulsar is subject to bias, because the distance prior to pulsars is not constant. Application to pulsar PSR J0218+4232 of the correct method for a realistic spatial distribution of millisecond pulsars shows that the isotropic gamma-ray flux of this recycled pulsar is more than 10% of its spin-down luminosity.

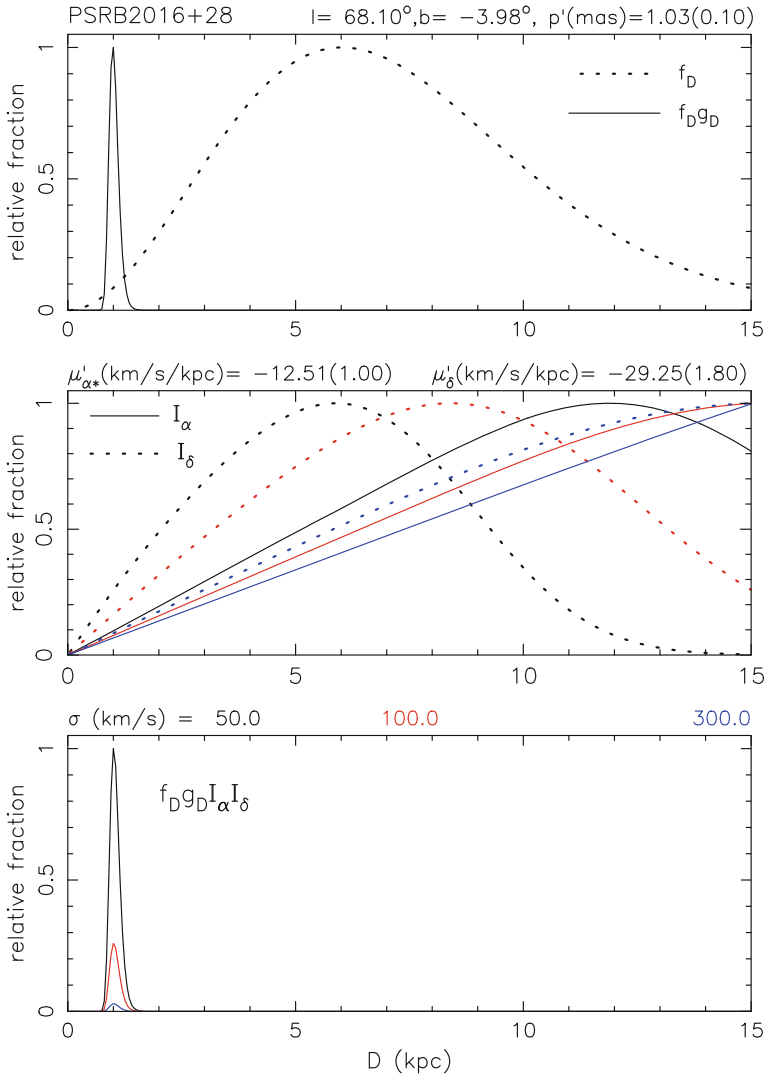


**Figure 8.** The integrand of equation (16), as in Fig. 6, but now for PSR B2016+28. In this case, the parallax and galactic pulsars distribution (*top*) and proper motion (*middle*) imply a small projected velocity $v'_\perp$. This favours the Maxwellian with low scaling parameter $\sigma = 50$ km/s.

For the determination of spatial velocities of young, in the sense of not recycled pulsars we only have measurements of the projections $v_\perp$ of these velocities on the celestial sphere. The most direct measurements of $v_\perp$ are obtained from VLBI observations of parallax and proper motion. Timing observations can also be used, but the measurement uncertainties are generally several orders of magnitude larger, allowing for determinations of proper motions, but only giving upper limits to the parallaxes. Indirect measurements of distances from dispersion measures depend on models for the electron distribution in the Milky Way, and as a result the uncertainties in the distances thus derived are large, and not gaussian but systematic.

Detailed analysis of the parallaxes and proper motions of 28 pulsars confirms the suspicion based on a rough analysis that a single Maxwellian does not describe the velocity distribution of these pulsars. A description with two Maxwellians is significantly better, and finds as a best solution that 42% of the pulsars follow a Maxwellian with distribution parameter $\sigma_1 = 77$ km/s and 58% a Maxwellian with $\sigma_2 = 320$ km/s. This detailed analysis considers pulsar velocities with respect to their local standard of rest, and to do so applies corrections for galactic rotation. At the current level of accuracy, however, it turns out that these corrections do not have a significant impact on the result.

The number of 28 pulsars for which accurate measurements are available is too small to conclude that the velocity distribution is indeed given by the sum of two Maxwellians. It is clear that pulsars have a wide

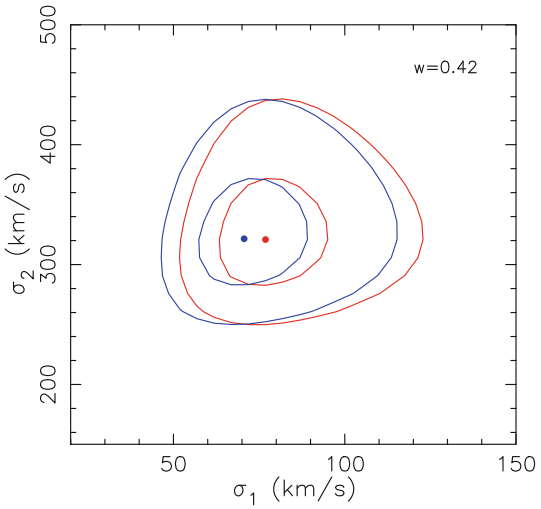


**Figure 9.** Red: Contours indicating the allowable range of the best solution for two Maxwellians, for $w = 0.42$. The best solution is given as a point, the contours contain 68% and 95% probability ($\Delta\mathcal{L}_{2\mathrm{maxw}} = 1$ and $\Delta\mathcal{L}_{2\mathrm{maxw}} = 4$, respectively). Blue: the same for a model in which the corrections for galactic rotation are omitted.

range of velocities, but to determine the exact form of the distribution, accurate measurements of more pulsars are necessary.

## Acknowledgements

The authors would like to thank Andrei Igoshev for discussions.

## Appendix

### *Appendix A. The Maxwellian velocity distribution and its projection*

The Maxwellian velocity distribution may be written as

$$f(v)\mathrm{d}v = \sqrt{\frac{2}{\pi}}\frac{v^2}{\sigma^3}\mathrm{e}^{-v^2/(2\sigma^2)}\mathrm{d}v. \tag{A.1}$$

In the isotropic case, the Maxwellian can be decomposed into three gaussian distributions with the same $\sigma$ but otherwise independent, along three mutually perpendicular directions. In the $x$-direction, for example, we have

$$f(v_x)\mathrm{d}v_x = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-{v_x}^2/(2\sigma^2)}\mathrm{d}v_x. \tag{A.2}$$

Choosing the $z$-direction along the line-of-sight, we find for the velocity perpendicular to the line-of-sight,

$$\begin{aligned} f(v_\perp)\mathrm{d}v_\perp &= \frac{1}{2\pi\sigma^2}\mathrm{e}^{-(v_x^2+v_y^2)/(2\sigma^2)}\mathrm{d}v_x\mathrm{d}v_y \\ &= \frac{1}{\sigma^2}\mathrm{e}^{-v_\perp^2/(2\sigma^2)}v_\perp\mathrm{d}v_\perp. \end{aligned} \tag{A.3}$$

The cumulative distribution of $v_\perp$ is as follows:

$$\begin{aligned} f(v_\perp < v_\mathrm{c}) &= \int_0^{v_c}\frac{1}{\sigma^2}\mathrm{e}^{-v_\perp^2/(2\sigma^2)}v_\perp\mathrm{d}v_\perp \\ &= 1 - \mathrm{e}^{-{v_c}^2/(2\sigma^2)}. \end{aligned} \tag{A.4}$$

## References

Bailer-Jones, C. A. L. 2015, *PASP*, **127**, 994.
Blaauw, A. 1985, in: *Birth and evolution of massive stars and stellar groups*, edited by W. Boland & H. van Woerden, Reidel, Dordrecht, p. 211.
Brisken, W. F., Benson, J. M., Goss, W. M., Thorsett, S. E. 2002, *Astrophys. J.*, **571**, 906.
Brown, A. G. A., Arenou, F., van Leeuwen, F., Lindegren, L., Luri, X. 1997, in: *ESA Symposium Hipparcos-Venice 1997*, *ESA SP* **402**, 63.
Chatterjee, S., Brisken, W. F., Vlemmings, W. H. T. *et al.* 2009, *Astrophys. J.*, **698**, 250.
Faucher-Giguère, C. -A., Kaspi, V. M. 2006, *Astrophys. J.*, **643**, 332.
Feast, M. 2002, *MNRAS*, **337**, 1035.
Francis, C. 2014, *MNRAS*, **444**, L6.
Hartman, J. W., Bhattacharya, D., Wijers, R., Verbunt, F. 1997, *Astron. Astrophys.*, **322**, 477.
Hobbs, G., Lorimer, D. R., Lyne, A. G., Kramer, M. 2005, *MNRAS*, **360**, 974.
Igoshev, A., Verbunt, F., Cator, E. 2016, *Astron. Astrophys.*, **591**, A123.
Kirsten, F., Vlemmings, W., Campbell, R. M., Kramer, M., Chatterjee, S. 2015, *Astron. Astrophys.*, **577**, A111.
Lutz, T. E., Kelker, D. H. 1973, *PASP*, **85**, 573.
Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling, W. T. 1986, *Numerical Recipes, The art of scientific computing*, Cambridge University Press, Cambridge, p. 289.
Sandage, A., Saha, A. 2002, *Astron. J.*, **123**, 2047.
Verbiest, J. P. W., Lorimer, D. R., McLaughlin, M. A. 2010, *MNRAS*, **405**, 564.
Verbiest, J. P. W., Weisberg, J. M., Chael, A. A., Lee, K. J., Lorimer, D. R. 2012, *Astrophys. J.*, **755**, 39
Verbiest, J. P. W., Lorimer, D. R. 2014, *MNRAS*, **444**, 1859.
Verbunt, F., Igoshev, A., Cator, E. 2017, *Astron. Astrophys.*, submitted.