



Extending the Event-weighted Pulsation Search to Very Faint Gamma-ray Sources

P. Bruel

► To cite this version:

P. Bruel. Extending the Event-weighted Pulsation Search to Very Faint Gamma-ray Sources. Astronomy and Astrophysics - A&A, 2019, 622, pp.A108. 10.1051/0004-6361/201834555 . hal-01975022

HAL Id: hal-01975022

<https://hal.science/hal-01975022>

Submitted on 24 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Extending the event-weighted pulsation search to very faint gamma-ray sources

P. Bruel

Laboratoire Leprince-Ringuet, Ecole polytechnique, CNRS/IN2P3, 91128 Palaiseau, France
e-mail: philippe.bruel@llr.in2p3.fr

Received 1 November 2018 / Accepted 17 December 2018

ABSTRACT

Context. Because of the relatively broad angular resolution of current gamma-ray instruments in the MeV–GeV energy range, the photons of a given source are mixed with those coming from nearby sources or diffuse background. This source confusion seriously hampers the search for pulsation from faint sources.

Aims. Statistical tests for pulsation can be made significantly more sensitive when the probability that a photon comes from the pulsar is used as a weight. However, computing this probability requires knowledge of the spectral model of all sources in the region of interest, including the pulsar itself. This is not possible for very faint pulsars that are not detected as gamma-ray sources or whose spectrum is not measured precisely enough. Extending the event-weighted pulsation search to such very faint gamma-ray sources would allow improving our knowledge of the gamma-ray pulsar population.

Methods. We present two methods that overcome this limitation by scanning the spectral parameter space, while minimizing the number of trials. The first one approximates the source to background ratio yielding a simple estimate of the weight while the second one makes use of the full spatial and spectral information of the region of interest around the pulsar.

Results. We tested these new methods on a sample of 144 gamma-ray pulsars already detected by the *Fermi* Large Area Telescope data. Both methods detect pulsation from all pulsars of the sample, including the ones for which no significant phase-averaged gamma-ray emission is detected.

Key words. gamma rays: general – pulsars: general – methods: data analysis – methods: statistical

1. Introduction

One of the important contributions of the *Fermi* Large Area Telescope (LAT; Atwood et al. 2009) to gamma-ray astronomy is the detection of many pulsars. Prior to the *Fermi* launch in 2008, at most ten gamma-ray pulsars were known (Thompson 2008). Ten years later, more than 200 have been detected¹, representing the most abundant Galactic source class at GeV energies (Acero et al. 2015). Detecting more pulsars, especially faint ones, is very important to characterizing as well as possible the gamma-ray pulsar population.

One way of finding gamma-ray pulsars is to search for periodic emission after phase-folding gamma-ray data with rotation ephemerides, obtained from contemporaneous timing observations conducted in radio or X-rays. Prior to *Fermi*'s launch, a pulsar timing consortium (PTC) was organized to support LAT observations of pulsars by providing accurate pulsar ephemerides (Smith et al. 2008). Another way of finding gamma-ray pulsars is to look at LAT catalog unassociated sources with a pulsar-like spectrum (see e.g. Clark et al. 2017; Pleunis et al. 2017) but, by construction, it does not target very faint gamma-ray sources. We thus focus on the former approach in the rest of this paper.

The sensitivity to pulsation depends on the purity of the event sample, which depends on the level of background and on the instrument performance, especially its angular resolution.

In the MeV–GeV energy range, the Galactic diffuse emission is bright and is the main source of background for candidate pulsars near the Galactic plane. Neighbouring sources can also significantly increase the level of background because of the finite angular resolution of the instrument. The LAT point spread function (PSF) at low energy is driven by multiple scattering of the electron and positron created by pair-conversion of the photon in the tracker. The 68% containment angle ranges from 5.5° at 100 MeV to 0.1° at 30 GeV (Atwood et al. 2009).

As shown in Bickel et al. (2008) and Kerr (2011), it is possible to efficiently mitigate the background problem by weighting the pulsation statistical test with the probability of each event to come from the pulsar position. To compute the exact probability, one needs a complete description of the region of interest (RoI) around the pulsar (large enough to fully take into account the effects of the PSF), including the spectral model of all gamma-ray sources in the RoI. So one prerequisite is to be able to measure the spectrum of the candidate pulsar, which is usually done by performing a maximum-likelihood fit of the RoI. When the fit leads to a clear detection of the pulsar, its spectrum is well measured and the photon probabilities can be computed. On the contrary, a weak detection of the pulsar or obviously its non detection precludes a good measurement of its spectrum. As a consequence, current event-weighted methods to search for pulsation cannot be applied to very faint gamma-ray sources.

The only possible way to overcome this limitation is to scan over the pulsar spectral parameters in order to find the ones maximizing the statistical test used to search for periodic emission. The obvious drawback is that the resulting significance must be

¹ <https://confluence.slac.stanford.edu/display/GLAMCOG/Public+List+of+LAT+Detected+Gamma-Ray+Pulsars>

corrected for the number of trials performed during the scan. Consequently, the challenge of extending the event-weighted methods to very faint pulsars consists of minimizing the number of trials when scanning the spectral parameter space, while maximizing the pulsation sensitivity, that is, getting as close as possible to the spectral parameters yielding the optimal weights.

After presenting in Sect. 2 the statistical test we used for pulsation searches, we present in Sect. 3 a method using a simple weight definition that does not involve a detailed spatial nor spectral modelling of the RoI. In Sect. 4, we describe a more complex method that takes advantage of the complete knowledge of the RoI and the full capabilities of the instrument.

Both methods are tested on a sample of 144 known LAT pulsars. This sample comprises the 117 pulsars from the LAT Second Gamma-ray Pulsar Catalog (2PC; [Abdo et al. 2013](#)), among which only two are not significant gamma-ray sources when analysing eight years of LAT data. Adding 27 post-2PC detected pulsars ([Hou & Smith 2014](#); [Laffon et al. 2014](#); [Smith et al. 2017](#)) to our sample increases the number of faint pulsars to 12. The results, based on the analysis of 8 year of Pass 8 SOURCE class data ([Atwood et al. 2013](#); [Bruehl et al. 2018](#)), are presented and discussed in Sect. 5.

2. The event-weighted H-test

To test the signal periodicity, we use the weighted H-test derived by [Kerr \(2011\)](#) from the original H-test proposed by [de Jager et al. \(1989\)](#):

$$H_{mw} = \max \left[Z_{mw}^2 - c \times (i - 1) \right], \quad 1 \leq i \leq m \quad (1)$$

The weighted H-test definition involves the weighted Z_{mw}^2 test:

$$Z_{mw}^2 = \frac{2}{\sum_{i=1}^N w_i^2} \times \sum_{k=1}^m (\alpha_{wk}^2 + \beta_{wk}^2), \quad (2)$$

where

$$\alpha_{wk} = \sum_{i=1}^N w_i \cos(2\pi k \phi_i), \quad (3)$$

and

$$\beta_{wk} = \sum_{i=1}^N w_i \sin(2\pi k \phi_i), \quad (4)$$

where the sums run over the list of N photons with pulsar rotational phase ϕ_i and weight w_i .

In the following, we have adopted the H-test parameter recommended values $m = 20$ and $c = 4$ (since they provide an omnibus test) and write H_w instead of H_{20w} . We note that one property of Z_w^2 and H_w is that they are not sensitive to a global scaling of the photon weights.

The analytic, asymptotic null distribution of H_w has been derived by [Kerr \(2011\)](#). Because we want to use H_w even with small event samples, we derive this distribution directly with Monte Carlo simulations, allowing us to know, for any $x > 0$, the probability $P(H_w > x)$, that H_w evaluated on a non-periodic sample can fluctuate to values larger than x . The calibration procedure is described in Appendices A and B. We show that $P(H_w > x)$ can be parameterized as a function of the sum of the weights, computed under the prescription that the maximum weight is one.

Rather than using directly H_w^{meas} , the measured value of H_w , when searching for pulsation, we use instead $P_w = -\log_{10} P(H_w > H_w^{\text{meas}})$, that we name pulsation significance for simplicity's sake. We note that the relation between P_w and n_σ , the significance expressed in number of σ , is given by $P_w = -\log_{10} (1 - \text{erf}(n_\sigma / \sqrt{2}))$. The 3, 4, and 5σ levels correspond to $P_w \sim 2.57, 4.20$ and 6.24 , respectively.

3. Simple weights

Let us consider an RoI centered on a pulsar. The probability that a photon originates from the pulsar depends on the position and spectrum of all gamma-ray sources in the ROI and the response functions of the instrument. We assume that all sources are steady and that the pulsar spectrum does not depend on phase. For a photon at a position Ω with an energy E , the probability that it comes from the pulsar is the ratio of the pulsar differential rate, r_{psr} , and the total differential rate at this energy and position. Defining r_{bkg} as the sum of all but the pulsar differential rates, we have:

$$w(E, \Omega) = \frac{r_{\text{psr}}(E, \Omega)}{r_{\text{psr}}(E, \Omega) + r_{\text{bkg}}(E, \Omega)}. \quad (5)$$

3.1. Simple weight definition

Since we are interested in very faint pulsars, we assume that the pulsar differential rate is negligible compared to r_{bkg} . Let us also assume that the background only comprises an isotropic diffuse emission. The background differential rate thus depends only on the energy E and we have:

$$w(E, \Omega) = r_{\text{psr}}(E, \Omega) / r_{\text{bkg}}(E). \quad (6)$$

We note that the resulting weight w is directly proportional to the pulsar absolute flux. Since the event-weighted statistical tests are unchanged when all weights are scaled by a constant, in the limit of very faint pulsars only the shape of the pulsar spectrum matters and not its absolute flux normalization. We also note that the weight position dependence only comes from the pulsar differential rate. Since the pulsar is a point source, the position dependence is described by the instrument PSF.

We can rewrite the weight as the product of two functions:

$$w(E, \Omega) = f(E) \times g(E, \Omega) \quad (7)$$

where $g(E, \Omega)$ contains all the position dependence and is defined such that it is 1 at the pulsar position at all energies. $f(E)$ is the weight at the pulsar position and depends on the pulsar and background spectral shapes and on the energy dependent part of the PSF normalization. The maximum of $f(E)$ is set to one to follow the maximum weight prescription mentioned in Sect. 2.

At a given energy, the LAT PSF is well approximated by a Moffat profile ([Moffat 1969](#))²:

$$K(x, s) = \frac{1}{4\pi s^2} \left(1 + \frac{x^2}{4s^2} \right)^{-2}, \quad (8)$$

where x is the angular distance to the source and s a scale parameter.

² Formerly referred to as King profile in https://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/Cicerone_LAT_IRFs/IRF_PSF.html

The integral $\int_0^\infty \int_0^{2\pi} K(x, s) dx d\theta = 1$ and we note that the integral up to a distance of $3s$ is about 0.68, so that the parameter s corresponds to a third of the PSF 68% containment angle. We can thus write the function $g(E, \Omega)$ as

$$g(E, x) = \left(1 + \frac{9x^2}{4\sigma_{\text{psf}}^2(E)} \right)^{-2}, \quad (9)$$

where $\sigma_{\text{psf}}(E)$ is the energy dependent PSF 68% containment angle, which for the LAT Pass 8 SOURCE event class can be parameterized as follows (Ackermann et al. 2013):

$$\sigma_{\text{psf}}(E) = p_0(E/100)^{p_1} \oplus p_2, \quad (10)$$

with $p_0 = 5.11^\circ$, $p_1 = -0.76$, $p_2 = 0.082^\circ$ and E in MeV and the addition is in quadrature.

The weight position-independent part, $f(E)$, corresponds to the weight at the position of the pulsar. Its computation is made simple by the fact that both the pulsar and background rates involve the same instrument effective area, which cancels out in Eq. (6). As a consequence $f(E)$ can be computed directly using the pulsar and background spectral shapes, as well as the energy dependent part of the PSF normalization ($1/4\pi s^2$ in Eq. (8)).

As shown in Abdo et al. (2013), the spectral shape of pulsars between 60 MeV and 60 GeV energy range can be modelled with a power law with an exponential cutoff:

$$dN/dE \propto (E/E_0)^{-\gamma} e^{-(E/E_c)^\beta}, \quad (11)$$

with, for most of pulsars, an index γ ranging from 0.5 to 2, an energy cutoff E_c between 0.6 and 6 GeV (and β set to 1 in 2PC). We consider three (γ, E_c) test-cases: (2.0, 0.6 GeV), (0.5, 0.6 GeV) and (2.0, 6 GeV). The latter two encompass the bulk of the pulsars, while the first one corresponds to pulsars whose pulsation is driven by low energy photons.

Regarding the background spectral shape, we use the spectrum of the Galactic diffuse emission towards the Galactic center. It can be modelled with a smoothly broken power law, with spectral indices of ~ 1.6 and ~ 2.5 below and above ~ 3 GeV, respectively³.

Figure 1 shows $f(E)$ for the three pulsar test-cases. Expressed as a function of $\log_{10} E$, the function f is Gaussian-like. The pulsar energy cutoff is responsible for the decrease of f at high energy. At low energy, its decrease is mainly driven by the increase of the instrument PSF. The center position parameter μ_w is located between 2.5 and 4.5 (corresponding to 30 MeV and 30 GeV, respectively) while the half 68% width parameter σ_w ranges from 0.3 to 0.45.

This result does not change qualitatively when using the spectrum of the Galactic diffuse emission at high Galactic latitude: for a given pulsar spectrum, changing the background spectral index moves the peak position of $f(E)$ but its shape remains Gaussian-like, with a width in the same $[0.3, 4.5]$ range. As a consequence, we can use the same weight definition for young pulsars, that are mostly near the Galactic ridge, and millisecond pulsars, that tend to lie at high Galactic latitude.

We approximate $f(E)$ with a Gaussian in $\log_{10} E$:

$$f(E) = e^{-0.5(\log_{10} E - \mu_w)^2 / \sigma_w^2}. \quad (12)$$

³ The spectral indices are estimated using the *Fermi*/LAT Galactic diffuse emission model `gll_iem_v06.fits` available at <https://fermi.gsfc.nasa.gov/ssc/data/access/lat/BackgroundModels.html>.

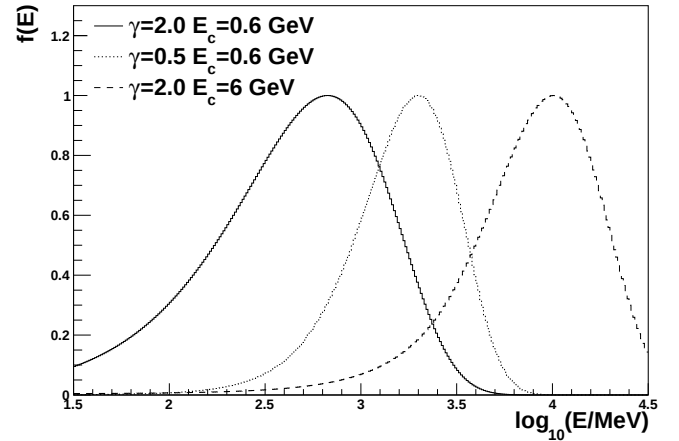


Fig. 1. Weight position independent part $f(E)$ for three test pulsar spectra.

We note that the examples in Fig. 1 are not exactly Gaussian. Using other functional forms that better model the negative tail of $f(E)$ does not improve the performance of the pulsation search, that is, better matching of $f(E)$ does not compensate for the simple assumptions that we use (very faint source on top of an isotropic background). Similarly, in Sect. 4.1 we show that refining the PSF information fails to improve the performance.

The Gaussian definition of $f(E)$ introduces two parameters, μ_w and σ_w . Formally, one would have to scan over these two parameters in order to fully explore the pulsar spectral parameter space. However, σ_w varies much less than μ_w , allowing us to fix σ_w and to vary only μ_w . We note that the σ_w range was obtained under the very faint pulsar hypothesis. For brighter pulsars, the energy range over which the pulsar emission is significant is larger, leading to a wider shape for $f(E)$, which corresponds to larger values of σ_w . For this reason, we chose to fix σ_w to 0.5. We describe our checks that, on average, this choice is valid for all pulsars in Sect. 5.

Thus, the simple weight definition that we propose is:

$$w(E, x, \mu_w) = e^{-2(\log_{10} E - \mu_w)^2} \left(1 + \frac{9x^2}{4\sigma_{\text{psf}}^2(E)} \right)^{-2}. \quad (13)$$

3.2. Simple weight scan

For a given value of μ_w , we can compute H_w and the corresponding pulsation significance P_w using the weights $w(E, x, \mu_w)$. Figure 2 shows P_w as a function of μ_w for PSR J1646–4346 (Smith et al. 2017). We chose this pulsar to illustrate the method because it is one of the pulsars in our sample that are not detected as a gamma-ray source with eight years of LAT data and, as a consequence, for which the original Kerr (2011) method can not be applied.

As can be seen in this example, $P_w(\mu_w)$ has a Gaussian shape around its maximum. This shape is simply the result of the relative matching between the weights $w(E, x, \mu_w)$ and the optimal weights, when μ_w varies. We refer to the maximum position and the 68% width of $P_w(\mu_w)$ as μ_P and σ_P , respectively. Figure 3 shows the distribution of σ_P versus μ_P for the 144 pulsar sample. The bulk of the pulsars have $3 < \mu_P < 4$ while a few have a low μ_P , corresponding to a pulsation signal driven by low energy photons (e.g. PSR J1513–5908, aka PSR B1509–58 (Abdo et al. 2010; Kuiper et al. 2017), whose

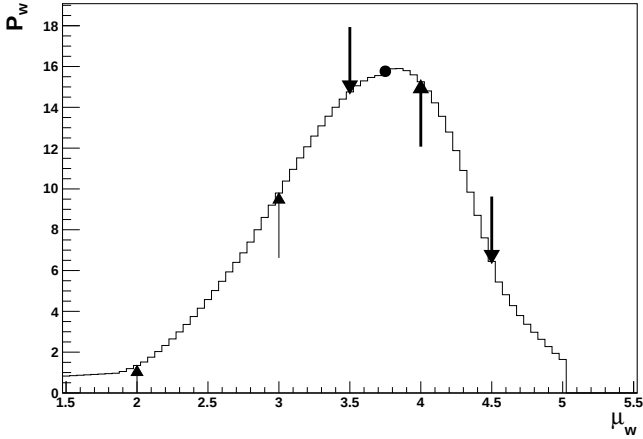


Fig. 2. Pulsation significance as a function of μ_w for PSR J1646–4346. The upward arrows correspond to the first three trials, while the downward arrows correspond to the two subsequent trials. The sixth trial, shown with a circle, is the maximum of the Gaussian passing through the three points corresponding to the thick-line arrows.

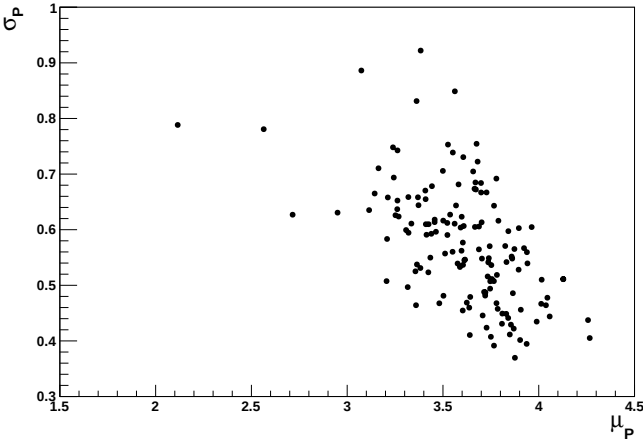


Fig. 3. 68% width as a function of the peak position of $P_w(\mu_w)$ for the 144 pulsar sample.

energy cutoff is below 30 MeV). The width σ_P is between 0.35 and 0.9.

Searching for pulsation consists in finding the maximum of $P_w(\mu_w)$. The easiest way to do so is to perform a very fine scan over μ_w but that would lead to a too large number of trials. Taking advantage of the Gaussian-shape of $P_w(\mu_w)$ around its maximum, we performed the following algorithm (illustrated in Fig. 2):

1. Test three values of $\mu_w = (2, 3, 4)$. Let μ_0 be the one giving the maximum P_w .
2. Test two more values of $\mu_w = (\mu_0 - 0.5, \mu_0 + 0.5)$. The 0.5 distance is chosen such that it is of the order of the minimum of the 68% width of $P_w(\mu_w)$. Let μ_1 be the one giving the maximum of P_w among all the trials with $2 \leq \mu \leq 4$;
3. Test $\mu_w = \mu_g$, the position of the maximum of the Gaussian passing through the three points $\mu_1 - 0.5, \mu_1, \mu_1 + 0.5$ (all tested in the previous steps);

The resulting pulsation significance of the simple method, corrected for the 6 trials, is:

$$P_s = \max_{\mu_w \text{ scan}} P_w - \log_{10} 6. \quad (14)$$

We note that the six trials are partially correlated, especially the ones with a small difference in μ_w . Unfortunately there is no simple way to estimate and take into account these correlations when correcting the pulsation significance and so we simply ignored them and adopted the conservative definition of P_s given by Eq. (14).

4. Model weights

To increase the sensitivity of the pulsation search, one has to go beyond the simple weight definition of Eq. (13). There are two possible approaches: (1) use the spatial and spectral description of all sources in the RoI; (2) take into account the full PSF information. While the former requires precise RoI modelling including spectral fits, the latter can be achieved by a minimal change of the weight definition, as it is a consequence of a refinement of the PSF definition.

4.1. PSF event types

One of the new features of Pass 8 (Atwood et al. 2013), the latest version of the reconstruction and selection of LAT data, is the introduction of the PSF event types. Prior to Pass 8, the events of a given class (e.g. SOURCE class) were divided into two subclasses, FRONT and BACK, depending on the location of the photon pair-conversion in the LAT tracker. Because of the different thicknesses of the converters in these two sections of the tracker, the PSF of the two subclasses are significantly different: at 1 GeV the FRONT and BACK PSF 68% containment angles are 1.2 and 0.61°, respectively. In Pass 8, thanks to a multivariate analysis, the events are divided into four event types PSF0, PSF1, PSF2 and PSF3, ordered in decreasing quality of the reconstructed direction, with a PSF 68% containment angle at 1 GeV of 1.8, 1.0, 0.66 and 0.42°, respectively.

It is straightforward to take into account the PSF events types in the context of the simple weight definition of Eq. (13): one just has to properly modify the $f(E)$ part of the weight induced by the $1/4\pi s^2$ factor of the normalization of the Moffat profile (Eq. (8)). We note that this modification is made simple by both the faint-source and uniform-background hypotheses.

Tested on LAT data, this modification failed to clearly increase the pulsation significance of known pulsars. This failure can be explained by two reasons. First, as soon as the faint-source hypothesis does not hold, the correction induced by the $1/4\pi s^2$ factor of the Moffat profile normalization is not correct: in the extreme situation in which the background is negligible, all photons should have a weight of approximately one, regardless of their PSF event type.

Moreover, taking into account the better PSF description given by the PSF event types is in fact equivalent to a better spatial description of the RoI. This improvement might not be large enough to compensate the crude approximation of the uniform-background hypothesis. In other words, one has to use a precise spatial and spectral description of the RoI to take full advantage of the additional information provided by the PSF event types. As a consequence, going beyond the simple weights presented in the previous section requires the full spatial and spectral description of the RoI around the pulsar.

4.2. Pulsar RoI description

To compute the weights for each pulsar, we performed a binned maximum-likelihood fit, following the standard *Fermi*/LAT

analysis procedure, of the $10.05 \times 10.05^\circ$ RoI centered on the pulsar, with a pixel size of 0.05° . We used 37 bins in $\log_{10} E$, between 63.1 MeV and 316.2 GeV.

We included in the source model the following components:

- the Galactic diffuse emission and the isotropic template (that accounts for the isotropic diffuse emission as well as the residual background)⁴;
- all point-like and extended sources from the preliminary eight-year *Fermi*/LAT source list (FL8Y)⁵, within five degrees of the RoI border.

If the closest point source to the pulsar is within a distance of 1.5 times the source 95% error radius, we considered the source to be the pulsar and set its position to the pulsar position. Otherwise we added a new source at the pulsar position.

We performed the maximum-likelihood fit with SOURCE class events (combining all PSF event types) above 160 MeV whose zenith angle is less than 90° to avoid Earth limb contamination, collected between 2008 August 4 and 2016 August 2. The significance of each source in the model is estimated using the test statistic, TS, defined as twice the difference in log-likelihood obtained with and without the source. A TS = 25 corresponds to $\sim 4\sigma$ significance (Mattox et al. 1996). The 160 MeV energy threshold for the maximum-likelihood fit has been chosen to mitigate the effect of systematic errors due to our imperfect knowledge of the Galactic diffuse emission.

Using the model derived by the fit, we built for each PSF event type the 3D map (sky position and energy) of the number of predicted events coming from the pulsar as well as the 3D map of the total number of predicted events. We were then able to compute the weights by simply dividing the former map by the latter map.

When computing H_w , we assigned to each event the weight of the bin corresponding to the event position and energy of the 3D map corresponding to the PSF event type of the event. In Kerr (2011), the weights are computed using the *Fermi*/LAT Science Tool *gtsrprob*⁶ that assigns to each event the probability that the event belongs to a given source of the RoI. Our binned approach is less CPU intensive and the chosen binning is fine enough to not induce any significant loss of sensitivity.

4.3. Spectral parameter scan

For bright gamma-ray pulsars, the maximum-likelihood fit gives a large TS for the pulsar and its spectral parameters are well estimated. We were thus able to use these spectral parameters to compute the weights, as done in Kerr (2011). On the contrary, for faint pulsars (TS < 25) or pulsars just above the TS = 25 threshold for which the spectrum is not precisely estimated, we have to scan over the spectral parameters.

Instead of using Eq. (11) to model the power law with an exponential cutoff, we used the following expression:

$$dN/dE \propto (E/E_0)^{-\gamma} e^{a(E_0^\beta - E^\beta)}, \quad (15)$$

as used in FL8Y and the forthcoming 4FGL catalog, with β given by the fit for bright pulsars or fixed to 0.667 for faint or not-detected ones. The formal change between the energy cut off and

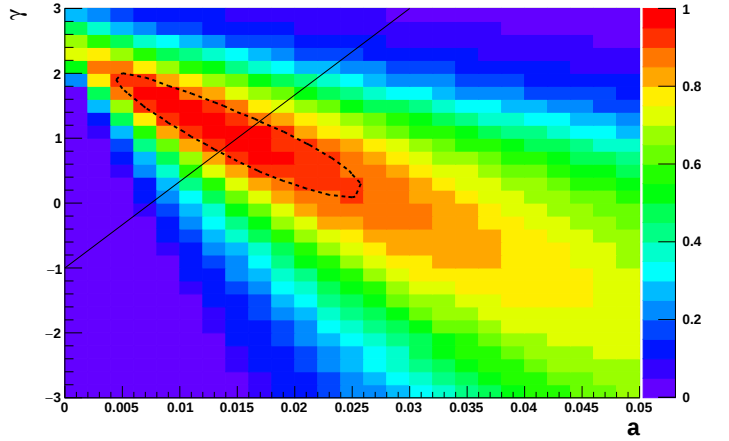


Fig. 4. Pulsation significance in the (a, γ) parameter space normalized relative to its maximum for PSR J1646–4346. The dashed contour line corresponds to the 90% level. The solid straight line corresponds to the line L_m along which a scan is performed to obtain the model weight pulsation significance, as described in Sect. 4.4.

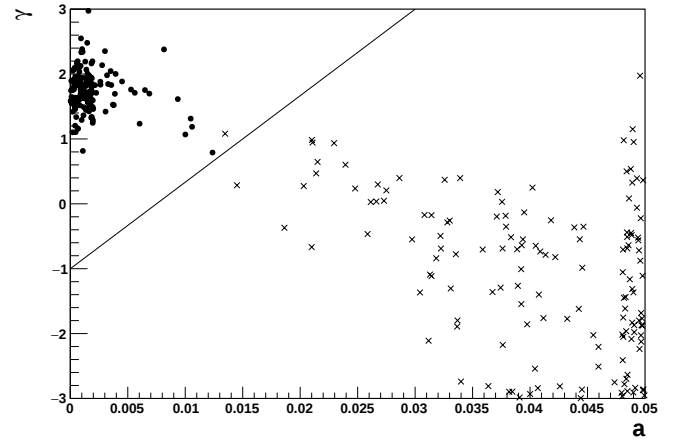


Fig. 5. Position of the lowest (dots) and highest (crosses) a points of E_{90} (the region of the (a, γ) plane with P_w larger than 90% of the maximum) for the 144 pulsar sample. The solid line corresponds to the line L_m along which a scan is performed to obtain the model weight pulsation significance, as described in Sect. 4.4.

parameter a is convenient to scan the spectral parameter space, as will be shown later.

For a given set of (a, γ) , we fix the corresponding spectral parameters and perform the maximum-likelihood fit with the pulsar normalization being the only free parameter. If TS < 4, we set the normalization to the 68% confidence limit. Using the resulting pulsar spectrum, we compute the four PSF event type weight maps and compute the pulsation significance, taking into account the PSF event type information by using for each event the PSF event type weight map corresponding to the event. Figure 4 shows how the pulsation significance varies in the (a, γ) plane for PSR J1646–4346.

The iso-level contours are elliptical. Let us consider the high P_w region corresponding to P_w larger than 90% of the maximum, whose ellipse-like contour is named E_{90} . To characterize E_{90} , we use the lowest and highest a extremities of E_{90} . These points are shown in Fig. 5 for the 144 pulsar sample. The important result is that the lowest and highest a points of E_{90} are not mixed and it is possible to define a line separating them.

⁴ <https://fermi.gsfc.nasa.gov/ssc/data/access/lat/BackgroundModels.html>

⁵ <https://fermi.gsfc.nasa.gov/ssc/data/access/lat/fl8y/>

⁶ <https://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/overview.html>

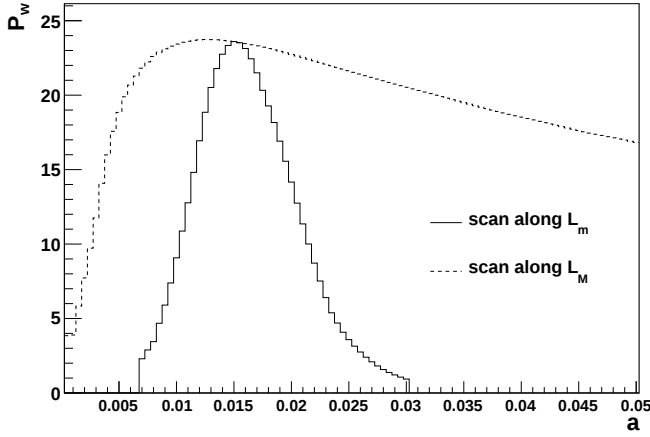


Fig. 6. Pulsation significance as a function of a , when scanning along L_m (solid line) and along L_M (dashed line) for PSR J1646–4346. L_m and L_M are two straight line in the pulsar parameter space (a, γ) . L_m goes from $(a = 0, \gamma = -1)$ to $(a = 0.03, \gamma = 3)$ while L_M approximates the $\mathcal{E}_{90}$ major-axis.

The fact that $\mathcal{E}_{90}$ is clearly elongated along a direction that tends to go from the upper-left to the lower-right of the (a, γ) plane is the consequence of how the weights vary with the spectral parameters. For a given background spectrum, when a increases, the effective position of the spectrum cutoff decreases and the energy position of the maximum weight at the pulsar position shifts to lower energy. On the other hand, the maximum weight energy position shifts to higher energy when γ decreases. So when a increases and γ decreases, the two effects partly counterbalance each other and the maximum weight energy position does not vary very much. The situation is reversed along the minor axis of $\mathcal{E}_{90}$: the increases of a and γ both shift the maximum weight energy position to lower energy. This is why $\mathcal{E}_{90}$ is very eccentric. We note that the convenient elliptical shape of the high P_w region is the result of the change from Eq. (11) to (15) to model the power law with an exponential cutoff.

4.4. Model weight scan

As for the simple weight method, the goal is to find the maximum of P_w in the lowest possible number of trials. The characterization of the high P_w region obtained in the previous section naturally suggests a two-step procedure: first finding the major-axis of $\mathcal{E}_{90}$ then scanning along the major-axis to find the maximum P_w .

To find the major-axis of $\mathcal{E}_{90}$ we performed a scan along the line L_m , going from $(a = 0, \gamma = -1)$ to $(a = 0.03, \gamma = 3)$, that crosses almost all $\mathcal{E}_{90}$, as shown in Fig. 5. The choice of L_m is motivated by the fact that increasing both a and γ at the same time allows a fast variation of the maximum weight energy position and, therefore, an efficient exploration of the pulsar spectrum parameter space. This choice might be slightly refined in the future thanks to the analysis of a larger sample of gamma-ray pulsars.

An example of the L_m scan is shown in Fig. 6. Because the P_w variation along L_m is Gaussian-like around the maximum, we can perform a similar six-trial algorithm as for the simple weight method, with the following test positions:

1. Test three values of $a = (0.005, 0.015, 0.025)$. Let a_0 be the one giving the maximum P_w .
2. Test two more values of $a = (a_0 - 0.005, a_0 + 0.005)$. The 0.005 distance is chosen such that it is of the order of the

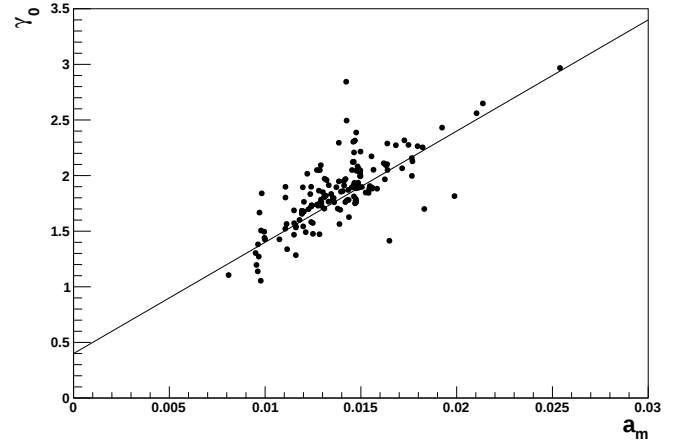


Fig. 7. Correlation between γ_0 , the intercept of the $\mathcal{E}_{90}$ major axis with the $a = 0$ axis, and a_m , the position of the maximum P_w along L_m for the 144 pulsar sample. The solid line shows the $\gamma_0 = 0.4 + 100a_m$ parameterization.

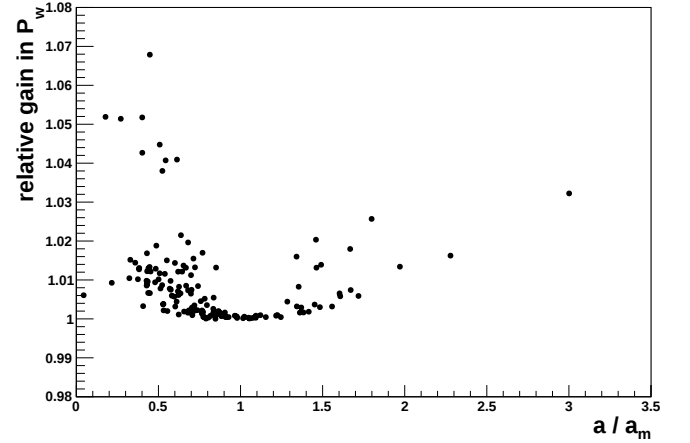


Fig. 8. Maximum P_w gain along L_M as a function of its relative position with respect to a_m for the 144 pulsar sample.

minimum of the 68% width of $P_w(a)$ when scanning along L_m . Let a_1 be the one giving the maximum of P_w among all the trials with $0.005 \leq a \leq 0.025$;

3. Test $a = a_g$ the position of the maximum of the Gaussian passing through the three points $a_1 - 0.005, a_1, a_1 + 0.005$ (all tested in the previous steps);
4. Let $A_m(a_m, \gamma_m)$ be the point giving the maximum P_w among the six trials.

We expect A_m to lie on the major-axis and we need one more point to define it. Figure 7 shows γ_0 , the intercept of the major-axis with the γ -axis, as a function of a_m . The correlation between the two can be modelled with $\gamma_0 = 0.4 + 100a_m$. We used this correlation to choose the point $A_0(a, \gamma) = (0, 0.4 + 100a_m)$ that, together with A_m , defines the major axis L_M .

An example of the P_w variation along L_M is shown in Fig. 6. The L_M scan allows us to reach a higher P_w than the L_m scan, but, because the latter goes through $\mathcal{E}_{90}$, the gain in P_w is rather modest (less than 10% by definition of $\mathcal{E}_{90}$). To optimize the definition of the second step, we look at the maximum relative gain in P_w along L_M versus its relative position with respect to a_m . As can be seen in Fig. 8, the gain in P_w is on average about 1%, reaching at most 5% for few pulsars. Regarding the optimal position a , it can be either smaller or larger than a_m . This would

imply at least two more trials, on top of the six ones already performed during the L_m scan, inducing an additional trial correction of $\log_{10}(8/6) = 0.125$.

For pulsars on the verge of pulsation detection, that is, at the 4σ level, corresponding to $P_w \sim 4.2$, the additional trial correction corresponds to about 2% of P_w , larger than the average 1% gain of the L_M scan. As a consequence, we chose not to perform the L_M scan. The model weight pulsation significance corrected for the six trials is thus:

$$P_m = \max_{L_m \text{ scan}} P_w - \log_{10} 6. \quad (16)$$

As in the simple weight method, the six trials are partially correlated and, for the same reasons, we choose to ignore them and to use a conservative definition of P_m .

5. Results

To test the performance of the simple and model weight methods, we applied them to the sample of 144 LAT pulsars (117 pulsars from 2PC and 27 post-2PC detected pulsars). For each pulsar, we performed the binned maximum-likelihood fit presented in Sect. 4.2 to estimate the TS of the pulsar and computed the model weights. 12 pulsars are found to have $TS < 25$.

To perform the pulsation search, we selected Pass 8 SOURCE class events within five degrees of the pulsar above 60 MeV. We used the *Fermi* plugin (Ray et al. 2011) to the pulsar timing software TEMPO2 (Hobbs et al. 2006; Edwards et al. 2006) and the pulsar ephemeris provided by the PTC (Smith et al. 2008, 2019) to convert the event arrival time into a pulsar rotational phase. The pulsation probabilities are computed using only the data collected during the validity period of the ephemerides.

We also compared the simple and model weight results to that obtained with the original Kerr (2011) method, which simply corresponds to the model weight method with the weights computed using the spectral parameters given by the maximum-likelihood fit. This is possible only when the source is significantly detected by the fit, meaning when $TS > 25$. We named the corresponding pulsation significance P_{fit} . Compared to P_m , P_{fit} has the obvious advantage of being the result of only one trial.

To compare the performance of the new methods to a standard unweighted pulsation search, we also performed a grid search, with each grid point corresponding to a set of cuts in energy and distance to the pulsar. We used the grid parameters defined by Kerr (2011), where it is reported that the event weighting method improves the pulsation sensitivity by a factor 1.5–2.

5.1. Model weights

The pulsation significance P_m obtained with the model weights is shown in Fig. 9 as a function of TS. All 144 pulsars have $P_m > 5.4$, corresponding to 4.6σ . The general trend is that P_m on average increases with TS. This is not surprising since, on average, the larger the TS, the stronger the pulsation signal. The scatter around this trend is mainly due to the pulse shape diversity. The most interesting result is the clear detection of pulsation from all pulsars with $TS \lesssim 50$, which proves that the model weight method is able to detect pulsation even when the pulsar spectrum information is not available or not fully reliable. The P_m and P_s values are given in Table 1 for the 12 pulsars with $TS < 25$. Ignoring the PSF event type information when

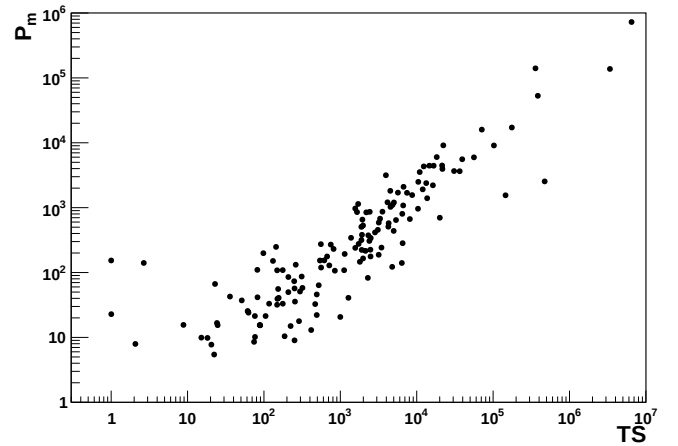


Fig. 9. Model weight pulsation significance P_m as a function of TS for the 144 pulsar sample. Pulsation is detected for all pulsars, including the pulsars with $TS < 25$, whose results are reported in Table 1.

Table 1. Results of the $TS < 25$ pulsars: TS of the spectral fit, model and simple weight pulsation significances, P_m and P_s .

Pulsar	(l, b)	TS	P_m	P_s
J0922+0638	(225.4, 36.4)	15.2	9.9	8.4
J1224–6407	(300.0, –1.4)	24.3	16.7	13.8
J1455–3330	(330.7, 22.6)	24.8	15.4	12.7
J1513–5908	(320.3, –1.2)	2.7	140.5	144.0
J1646–4346	(341.1, 1.0)	1.0	22.8	15.0
J1739–3023	(358.1, 0.3)	2.1	7.9	6.1
J1801–2451	(5.3, –0.9)	1.0	153.4	125.4
J1831–0952	(21.9, –0.1)	8.8	15.6	10.7
J1832–0836	(23.1, 0.3)	20.5	7.7	8.0
J1856+0113	(34.6, –0.5)	22.9	66.7	61.3
J1909–3744	(359.7, –19.6)	18.2	9.8	6.8
J2317+1439	(91.4, –42.4)	22.3	5.4	4.5

computing the model weights leads to a loss of sensitivity that decreases with TS, from 20% on average for faint pulsars to 10% for the brightest.

The comparison between P_m and P_{fit} is shown in Fig. 10. On average the model weight method provides the same pulsation significance, within $\sim 5\%$ for most of the pulsars. Only one has P_m lower than P_{fit} by more than 10%.

On the contrary, the model weight method often improves significantly over P_{fit} , with a gain greater than 20% for eight pulsars. These eight pulsars all have $TS > 60$ so it is unlikely that the improvement is explained by a poor estimation of the spectral parameters by the fit. A possible explanation is the presence of a significant off-pulse emission: the spectral parameters derived by the fit correspond to the spectrum of the sum of the pulsed and unpulsed gamma-ray components and not to the pulsed component alone, while the model weight method is able to find spectral parameters closer to the pulsed component spectrum.

As reported in Abdo et al. (2013), 34 2PC pulsars have a significant off-peak emission, whose spectrum is compatible with a simple power-law spectrum for 13 of them. Moreover, the off-peak emission of these 13 pulsars is generally soft, with a spectral index of approximately two. A flat and soft off-peak emission that is not negligible compared to the pulsed emission could lead to a total emission spectrum significantly different

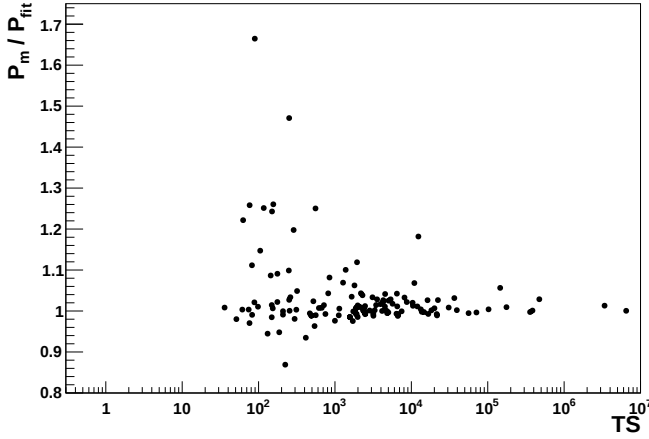


Fig. 10. Comparison of the model weight pulsation significance, P_m , and the pulsation significance derived when using the pulsar spectrum given by the spectral fit, P_{fit} , as a function of TS for the 131 pulsars with TS > 25 out of the 144 pulsar sample. The two methods give on average the same result but the model weight method is more sensitive by at least 20% for eight pulsars, whose results are reported in Table 2.

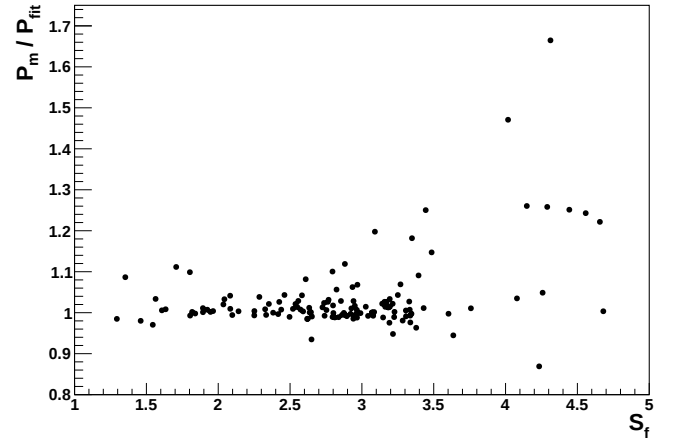


Fig. 11. Comparison of the model weight pulsation significance, P_m , and the pulsation significance derived when using the pulsar spectrum given by the spectral fit, P_{fit} , as a function of the “soft-and-flat” estimator S_f for the 131 pulsars with TS > 25 out of the 144 pulsar sample. The eight pulsars with $P_m/P_{\text{fit}} > 1.2$ are among the pulsars with the largest S_f , meaning those with the softest and flattest spectrum.

from the pulsed emission spectrum. Investigating that it is also the case of the largest P_m/P_{fit} pulsars would require performing the analysis of the off-pulse emission, which is out of the scope of this paper and will be presented in the forthcoming third LAT Gamma-ray Pulsar Catalog (3PC).

We instead estimated the significance of the curvature of the spectrum. As in [Abdo et al. \(2013\)](#) and [Acero et al. \(2015\)](#) we performed a maximum-likelihood fit assigning the pulsar a power-law spectrum and compute $\sigma_{\text{curv}} = (\text{TS} - \text{TS}_{\text{PL}})^{-1/2}$, where TS and TS_{PL} correspond to the maximum-likelihood fit of a power-law with and without exponential cutoff, respectively. We note however that σ_{curv} measures the significance of the curvature and not how flat or soft the spectrum is.

A very simple and crude estimator of the softness and flatness of the spectrum is provided by the sum $S = \gamma + \log_{10} E_c$, with $E_c = a^{-1/\beta}$. This estimator can be improved by taking into account the correlation between γ and $\log_{10} E_c$. The analysis of the distribution of γ vs $\log_{10} E_c$ for the 144 pulsar sample yields a positive correlation with a slope ~ 1.418 and the projection along the first principal axis is $S_f = 0.576\gamma + 0.817 \log_{10} E_c$. In the case of a spectrum without significant curvature, the spectral parameters γ and E_c are not precisely measured. To take into account this case, we replace γ and E_c by their 68% upper limit and define the following simple “soft-and-flat” estimator:

$$S_f = 0.576(\gamma + \delta\gamma) + 0.817 \log_{10}(E_c + \delta E_c). \quad (17)$$

Figure 11 shows the ratio P_m/P_{fit} as a function of S_f . Although the correlation between these two quantities is not perfect, we note that all the eight pulsars with $P_m/P_{\text{fit}} > 1.2$ lie in the right-hand tail of the S_f distributions, meaning that their spectra are among the most soft and flat of the sample. Some of their properties are given in Table 2. Only two of them have a curvature significance above 4σ : J0742–2822 and J1828–1101 with $\sigma_{\text{curv}} = 8$ and 4.9, respectively. All the others have $\gamma > 2.4$. We note that only four of these eight pulsars are in 2PC and only two are detected as a gamma-ray source (J0742–2822 and J1410–6132). Out of these two, only J1410–6132 is reported to have a significant off-peak emission.

The comparison of the model weights with the grid search is shown in Fig. 12 as a function of TS for the 144 pulsar sample.

Table 2. Pulsars with TS > 25 and $P_m/P_{\text{fit}} > 1.2$.

Pulsar	TS	P_m	P_{fit}	γ	S_f
J0729–1448	62.8	24.1	19.7	2.5 ± 0.1	4.7
J0742–2822 [†]	556.2	273.8	219.0	1.7 ± 0.2	3.4
J1151–6108	150.0	39.3	31.6	2.4 ± 0.1	4.6
J1410–6132	116.9	33.1	26.5	2.7 ± 0.1	4.4
J1431–4715	76.4	10.2	8.1	2.7 ± 0.1	4.3
J1531–5610	155.9	40.8	32.3	2.6 ± 0.1	4.1
J1828–1101 [†]	251.4	9.0	6.1	2.1 ± 0.2	4.0
J1935+2025	89.2	15.4	9.2	2.7 ± 0.1	4.3

Notes. The reported spectral index corresponds to the power-law fit, expect for the two $\sigma_{\text{curv}} > 4$ pulsars (indicated with a [†]), for which the spectral index is the result of the fit with the power law with an exponential cutoff.

We confirm the improvement of the weighting method over the unweighted approach, with a gain in sensitivity larger than two for the TS < 25 pulsars.

5.2. Simple weights

The comparison between the simple weights and the model weights is shown in Fig. 13. As expected, the simple weights are less powerful than the model weights. The difference in performance decreases from about 30% for the brightest pulsars to an average of about 15% at TS ~ 300 . This difference never goes beyond 40%. This difference is relatively small given the simplicity of the simple weight implementation compared to the complex procedure of the model weights. The P_s values are given in Table 1 for the 12 pulsars with TS < 25. Except for a few of the brightest pulsars, the simple model performs better than the standard grid search, especially for TS ≤ 100 where P_s is greater than P_{grid} by at least 40%.

When defining the simple weights in Sect. 3.1, we set $\sigma_w = 0.5$. To validate this choice, we perform a scan over σ_w between 0.2 and 0.95, with a 0.05 step. For each pulsar and each value of σ_w , we compute $r(\sigma_w)$, the ratio of P_s divided by the maximum P_s reached over the σ_w scan. The average over the 144 pulsar

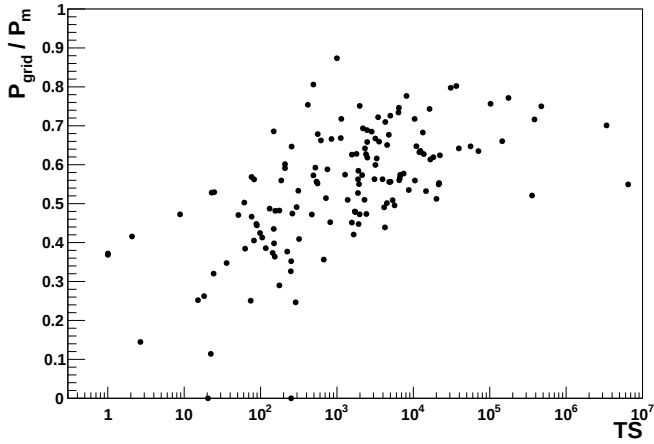


Fig. 12. Comparison of the pulsation significance obtained with an unweighted approach, P_{grid} , and the model weight pulsation significance, P_m , as a function of TS for the 144 pulsar sample. The improvement of the model weight method over the unweighted approach is large, especially for the lowest TS pulsars.

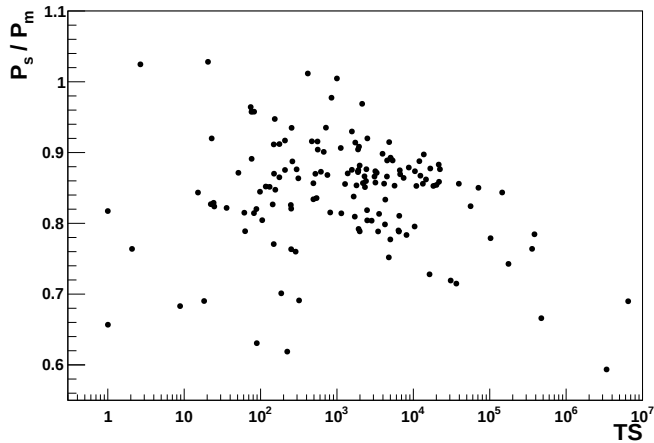


Fig. 13. Comparison of the simple and model weight pulsation significances as a function of TS for the 144 pulsar sample. The simple weight method is on average $\sim 15\%$ less sensitive than the model weight method.

sample of this ratio as a function of σ_w is shown in Fig. 14, as well as the lowest ratio, that is, corresponding to the pulsar for which the choice of σ_w is the worst. The two quantities have a maximum between 0.5 and 0.6 and do not vary much around the maximum, which validates the $\sigma_w = 0.5$ choice.

To summarize the results, we show in Fig. 15 the comparison of all methods that allows a clear ranking of the pulsation search methods, from the least sensitive un-weighted grid scan to the most powerful [Kerr \(2011\)](#) and model weight methods, with the latter, contrary to the former, being able to detect pulsation from very faint pulsars.

We note that both the simple and model weight approaches can be used at other wavelengths (e.g. X-ray and TeV bands) but they need to be adapted to the specific context of each wavelength, taking into account the pulsar spectral parameter phase space, the PSF energy dependence and the typical background spectrum. In the case of the simple weight method, the derived general shape of $f(E)$ may be very different from the Gaussian-like one obtained for the LAT energy band.

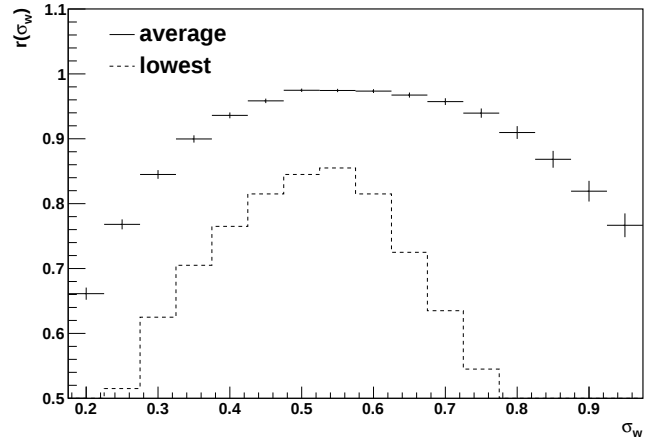


Fig. 14. Average (solid) and lowest (dashed) ratio $r(\sigma_w)$ over the 144 pulsar sample as a function of σ_w . For each σ_w , $r(\sigma_w)$ is the ratio of P_s divided by the maximum P_s reached over the σ_w scan.

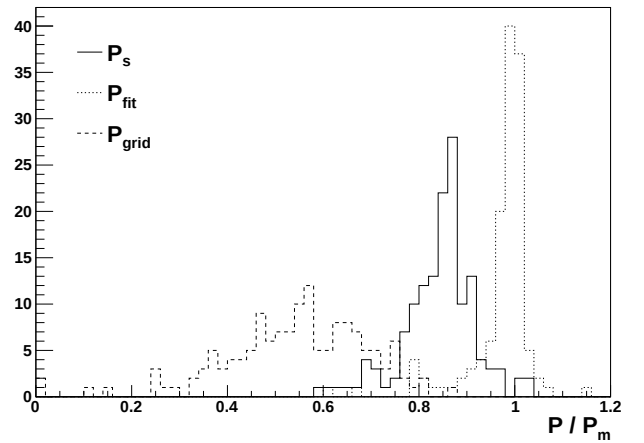


Fig. 15. Ratio of the pulsation significance with respect to P_m for the simple weights (solid), original event weight (dotted) and the grid search (dashed), for the 144 pulsar sample.

6. Conclusion

We show that it is possible to extend the event weighting technique to very faint gamma-ray sources when searching for pulsation and presented two approaches. The first one, the simple weight method, uses a very simple definition of the weights while the second one, the model weight method, fully takes into account the spatial and spectral information of the RoI around the pulsar. The key point for both methods is to explore efficiently the pulsar spectral parameter space, which is done with only six trials.

The model weight method reaches the same performance as the original event weighting method that was not applicable to the very faint gamma-ray sources. It can even be more powerful in the case of pulsars with a significant off-pulse emission.

The simple weight method is less sensitive than the model weight one but the loss of sensitivity is only $\sim 30\%$ for faint sources. So this simple approach can be very useful, especially since it is straightforward to implement and is much less CPU intensive.

As a consequence the model weights method is on average the most sensitive method. It works for all pulsars (faint or bright, with or without off-pulse emission) and naturally benefits from any improvement of the instrument performance (e.g. Pass 8 PSF event types).

After ten years in orbit, *Fermi*/LAT continues to take data that are folded with the updated ephemerides provided by the

Pulsar Timing Consortium. The two new methods presented in this paper are designed to help to detect new gamma-ray pulsars, including very faint ones, allowing us, for instance, to further test the existence of a spin-down power “deathline” (Smith et al. 2019) below which pulsars might cease to produce gamma-ray emission.

Acknowledgements. We thank our *Fermi*/LAT collaborators Toby Burnett, Lucas Guillemot, David Smith and Matthew Kerr for fruitful discussions. The Nançay Radio Observatory is operated by the Paris Observatory, associated with the French Centre National de la Recherche Scientifique (CNRS). The Parkes radio telescope is part of the Australia Telescope 533 which is funded by the Commonwealth Government for operation as a National Facility managed by CSIRO. We thank our colleagues for their assistance with the radio timing observations. The Lovell Telescope is owned and operated by the University of Manchester as part of the Jodrell Bank Centre for Astrophysics with support from the Science and Technology Facilities Council of the United Kingdom. The *Fermi*/LAT Collaboration acknowledges generous ongoing support from a number of agencies and institutes that have supported both the development and the operation of the LAT as well as scientific data analysis. These include the National Aeronautics and Space Administration and the Department of Energy in the United States, the Commissariat à l’Energie Atomique and the Centre National de la Recherche Scientifique / Institut National de Physique Nucléaire et de Physique des Particules in France, the Agenzia Spaziale Italiana and the Istituto Nazionale di Fisica Nucleare in Italy, the Ministry of Education, Culture, Sports, Science and Technology (MEXT), High Energy Accelerator Research Organization (KEK) and Japan Aerospace Exploration Agency (JAXA) in Japan, and the K. A. Wallenberg Foundation, the Swedish Research Council and the Swedish National Space Board in Sweden. Additional support for science analysis during the operations phase is gratefully acknowledged from the Istituto Nazionale di Astrofisica in Italy and the Centre National d’Études Spatiales in France. This work performed in part under DOE Contract DE-AC02-76SF00515.

References

- Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, *ApJ*, 714, 1
- Abdo, A. A., Ajello, M., Allafort, A., et al. 2013, *ApJS*, 208, 17
- Acero, F., Ackermann, M., Ajello, M., et al. 2015, *ApJS*, 218, 23
- Ackermann, M., Ajello, M., Allafort, A., et al. 2013, *ApJ*, 765, 1
- Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, *ApJ*, 697, 1071
- Atwood, W., Albert, A., Baldini, L., et al. 2013, *eConf C121028*, 8, in *Proc. 4th Fermi Symposium* (Monterey)
- Bickel, P., Kleinj, B., & Rice, J. 2008, *ApJ*, 685, 384
- Bruel, P., Burnett, T. H., & Digel, S. W. 2018, *ArXiv e-prints* [arXiv:1810.11394]
- Clark, C. J., Wu, J., Pletsch, H. J., et al. 2017, *ApJ*, 834, 106
- de Jager, O. C., & Büsching, I. 2010, *A&A*, 517, L9
- de Jager, O. C., Raubenheimer, B. C., & Swanepoel, J. W. H. 1989, *A&A*, 221, 180
- Edwards, R. T., Hobbs, G. B., & Manchester, R. N. 2006, *MNRAS*, 372, 1549
- Hobbs, G. B., Edwards, R. T., & Manchester, R. N. 2006, *MNRAS*, 369, 655
- Hou, X., Smith, D. A., et al. 2014, *A&A*, 570, A44
- Kerr, M. 2011, *ApJ*, 732, 38
- Kuiper, L., Hermsen, W., & Dekker, A. 2017, *MNRAS*, 475, 1
- Laffon, H., Smith, D. A., Guillemot, L., et al. 2014, in *Proc. 5th Fermi symposium* (Nagoya)
- Mattox, J. R., Bertsch, D. L., Chiang, J., et al. 1996, *ApJ*, 461, 396
- Moffat, A. F. J. 1969, *A&A*, 3, 455
- Pleunis, Z., Bassa, C. G., Hessels, J. W. T., et al. 2017, *ApJ*, 846, L19
- Ray, P. S., Kerr, M., Parent, D., et al. 2011, *ApJS*, 194, 17
- Smith, D. A., Guillemot, L., Camilo, F., et al. 2008, *A&A*, 492, 923
- Smith, D. A., Guillemot, L., Kerr, M., et al. 2017, in *Proc. 11th INTEGRAL conference* (Amsterdam)
- Smith, D. A., Bruel, P., Cognard, I., et al. 2019, *ApJ*, 871, 78
- Thompson, D. J. 2008, *Rep. Prog. Phys.*, 71, 116901

Appendix A: Monte Carlo estimated probability distribution of H_{20}

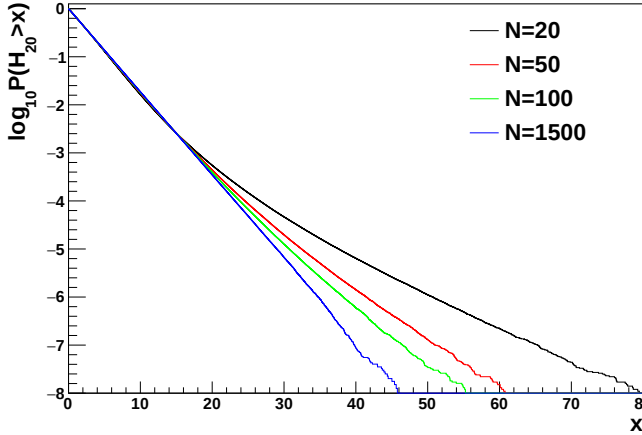


Fig. A.1. Logarithm of the H_{20} cumulative distribution for various data sample sizes.

The asymptotic H_{20} probability distribution has been estimated using Monte Carlo (MC) by [de Jager et al. \(1989\)](#) and [de Jager & Büsching \(2010\)](#). It corresponds to a simple exponential $P(H_{20} > x) = e^{-\lambda x}$ with $\lambda \sim 0.4$. This result is very close to the analytic solution found by [Kerr \(2011\)](#) that yields a practical formula (valid for $m \geq 10$) with $\lambda = 0.398405$.

Pulsation searches very often require multiple trials corresponding to different event selections. Some of these can lead to small data samples. As a consequence it is important to estimate the H_{20} probability distribution for low N , the number of events used when computing H_{20} . To do so, we ran 10^9 MC realizations for various N , assigning a random phase to each event. Figure A.1 shows the logarithm of the cumulative distribution, $\log_{10} P(H_{20} > x)$, for $N = 20, 50, 100$ and 1500 . The asymptotic behaviour is reached in $N = 1500$ case but the distribution of the other cases shows a clear departure from the pure exponential above $x \sim 20$.

To characterize this departure, we fitted $\log_{10} P(H_{20} > x)$ with a linear polynomial over the interval corresponding to $-7 < \log_{10} P(H_{20} > x) < -4$ (i.e. $x \gtrsim 23$). The slope λ_1 of this linear polynomial is shown in Fig. A.2 as a function of N . We find the following parameterization for λ_1 (shown in Fig. A.2):

$$\lambda_1(N) = \lambda_0 + 0.0525796e^{-N/215.170} + 0.086406e^{-N/35.5709} \quad (\text{A.1})$$

where $\lambda_0 = -0.398405/\log(10) = -0.173025$.

We first approximate $\log_{10} P(H_{20} > x)$ with a broken linear polynomial whose slopes at small and large x are λ_0 and $\lambda_1(N)$, respectively. We find that the break position is compatible with $x = 22$ for all N . For low values of N the broken linear polynomial approximation does not work well around the break position, as shown in Fig. A.3. To have a better representation of $\log_{10} P(H_{20} > x)$, we used a double broken linear polynomial approximation, with the same slopes at small and large x as for the single one, with break positions at $x = 15$ and $x = 29$:

$$\log_{10} P(H_{20} > x) \sim \begin{cases} \lambda_0 x & \text{if } x < 15, \\ 15\lambda_0 + \frac{\lambda_0 + \lambda_1(N)}{2}(x - 15) & \text{if } 15 < x < 29 \\ 22\lambda_0 + \lambda_1(N)(x - 22) & \text{if } x > 29. \end{cases} \quad (\text{A.2})$$

When considering the x interval such that $P(H_{20} > x) > 10^{-7}$, the maximum of the absolute difference between the

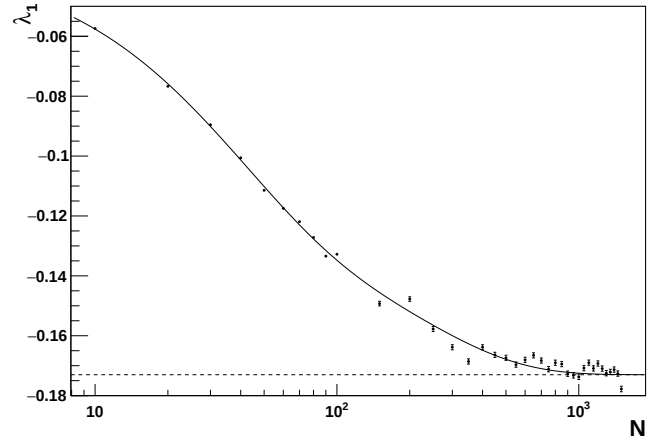


Fig. A.2. Parameterization of λ_1 , the slope of $\log_{10} P(H_{20} > x)$ over the interval corresponding to $-7 < \log_{10} P(H_{20} > x) < -4$, as a function of N .

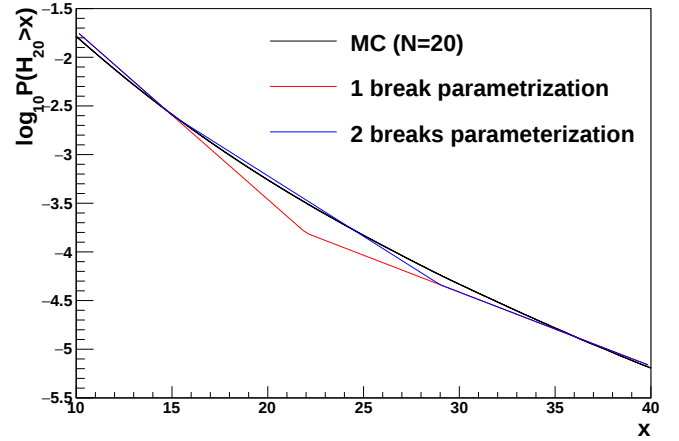


Fig. A.3. Logarithm of the H_{20} cumulative distribution for $N = 20$ in the x range where a single broken linear polynomial does not provide a good approximation. The black line corresponds to the MC result while the red and blue curves correspond to the single and double broken linear polynomial approximations, respectively.

parameterization and the MC result is less than 0.1 for $N \geq 20$. For $N = 10$, the maximum difference reaches 0.25 at $x = 29$. As a consequence, the approximation given by Eq. (A.2) is valid for $N \geq 20$ with a 0.1 precision for $P(H_{20} > x) > 10^{-7}$.

Appendix B: Monte Carlo estimated probability distribution of H_{20w}

To estimate the probability distribution of H_{20w} , we ran 10^8 MC realizations of a five-degree region of interest with a uniform background following the Galactic diffuse emission spectrum defined in Sect. 3.1 (corresponding to a broken power law, with spectral indices of ~ 1.6 and ~ 2.5 below and above ~ 3 GeV, respectively), convoluted with the *Fermi*/LAT effective area. These realizations are performed for various numbers of events. For each realization, we compute H_{20w} using the simple weights definition of Eq. (13) with σ_{psf} set to one degree independently of energy. For CPU efficiency's sake, each realization is used twice, with μ_w set to 2.5 and 3, each choice leading to a value of H_{20w} .

When parameterizing $\log_{10} P(H_{20} > x)$ in the previous Section, we used the number of events because it drives the level of fluctuations. In the case of the weighted version of H_{test} ,

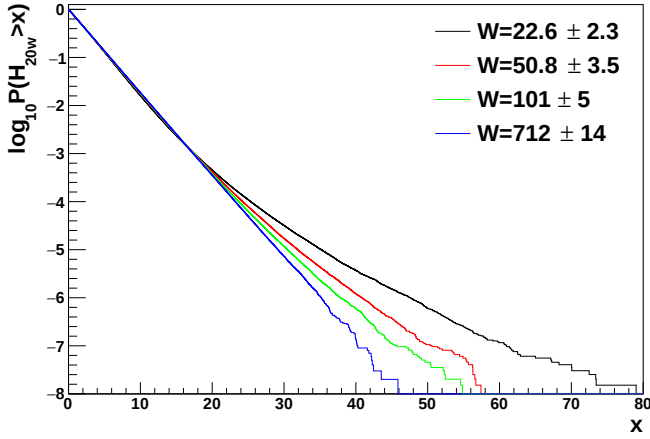


Fig. B.1. $\log_{10} P(H_{20w} > x)$ for various sum of weights.

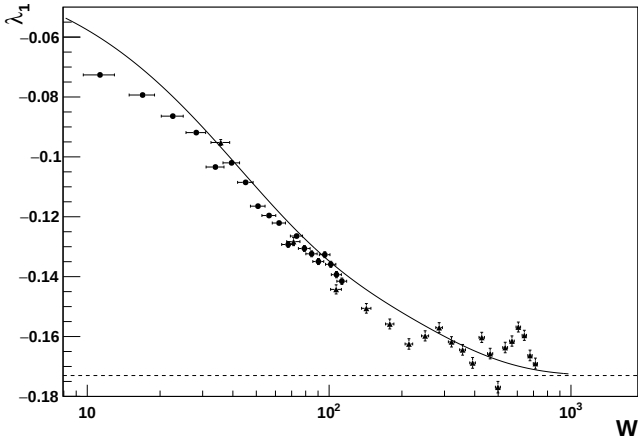


Fig. B.2. Slope of $\log_{10} P(H_{20w} > x)$ over the interval corresponding to $-7 < \log_{10} P(H_{20w} > x) < -4$ as a function of W . Circles and triangles correspond to $\mu_w = 2.5$ and 3 , respectively.

the number of events is not useful anymore. We used instead the sum of the weights, W , that we computed under the prescription that the maximum weight is one. Figure B.1 shows $\log_{10} P(H_{20w} > x)$ for some of the MC configurations corresponding to $W \sim 20, 50, 100, 700$. As in the unweighted case, $\log_{10} P(H_{20w} > x)$ departs from a pure exponential at low W and can be approximated to first order by a broken linear polynomial. We note that the various MC configurations (number of events,

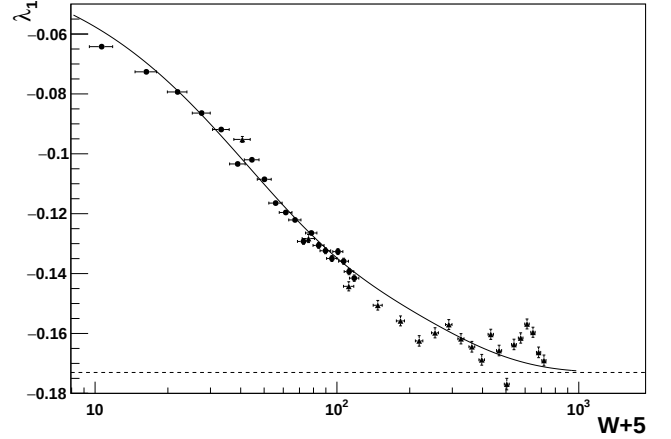


Fig. B.3. Slope of $\log_{10} P(H_{20w} > x)$ over the interval corresponding to $-7 < \log_{10} P(H_{20w} > x) < -4$ as a function of $W + 5$. Circles and triangles correspond to $\mu_w = 2.5$ and 3 , respectively.

μ_w choice) were chosen to explore a range for W between about 10 and 1000.

We estimate λ_1 , the slope over the interval corresponding to $-7 < \log_{10} P(H_{20w} > x) < -4$. Figure B.2 shows how the λ_1 variation with W compares to the unweighted-case λ_1 parameterization of Eq. (A.1). This parameterization works well for large values of W but not at low W . We find that using $W + 5$ instead of W gives a better match, as shown in Fig. B.3. The fact that increasing W leads to a better match is not surprising since, in the weighted case, the number of events that play a significant role in the H_{20w} computation (i.e. with a relatively large weight) is on average larger than W .

As a consequence, we used the following parameterization of $\log_{10} P(H_{20w} > x)$, that is obtained by simply replacing N by $W + 5$ in Eq. (A.2):

$$\log_{10} P(H_{20w} > x) \sim \begin{cases} \lambda_0 x & \text{if } x < 15, \\ 15\lambda_0 + \frac{\lambda_0 + \lambda_1(W+5)}{2}(x - 15) & \text{if } 15 < x < 29, \\ 22\lambda_0 + \lambda_1(W+5)(x - 22) & \text{if } x > 29. \end{cases} \quad (\text{B.1})$$

When considering the x interval with $P(H_{20w} > x) > 10^{-7}$, the maximum of the absolute difference between the parameterization and the MC result is less than 0.1 for $W \geq 10$. As a consequence, the approximation given by Eq. (B.1) is valid for $W \geq 10$ with a 0.1 precision for $P(H_{20w} > x) > 10^{-7}$.