A Stacking Survey of Gamma-ray Pulsars

Yuzhe Song(宋宇哲)1,2,3,4,5★ Timothy A. D. Paglione,1,2,3 Joshua Tan3,6, Charles Lee-Georgescu6 and Danisbel Herrera3,7

1Physics Program, the Graduate Center, City University of New York, 365 Fifth Ave., New York, NY 10016, USA
2Department of Earth & Physical Sciences, York College, City University of New York, 94-20 Guy R. Brewer Blvd., Jamaica, NY 11451, USA
3Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
4Center for Astrophysics and Supercomputing, Swinburne University of Technology, John St, Hawthorn, VIC 3122, Australia
5OzGrav, ARC Center for Excellence in Gravitational Wave Discovery
6Department of Natural Sciences, LaGuardia Community College, City University of New York, 31-10 Thomson Ave, Long Island City, NY 11101, USA
7Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA

Accepted 2023 July 17. Received 2023 July 15; in original form 2023 May 15

ABSTRACT
We report on a likelihood stacking search for gamma-ray pulsars at 362 high-latitude locations that coincide with known radio pulsar positions. We observe a stacked signal conservatively 2.5σ over the background. Stacking their likelihood profiles in spectral parameter space implies a pulsar-like spectral index and a characteristic flux a factor of 2 below the Fermi-LAT point source sensitivity, assuming isotropic/unbeamed emission from all sample pulsars. The same procedures performed on empty control fields indicate that the pulsars as a population can be distinguished from the background with a Δ(TS) = 28 at the peak location (or 4.8σ), and the stacked spectra of the control fields are distinctly softer than those of the pulsars. This study also probes a unique region of parameter space populated by low E pulsars, most of which have no γ-ray ephemeris available, and is sensitive to high duty cycles. We also discusses the possible γ-ray emission mechanism from such pulsars.

Key words: pulsars: general – stars: neutron – gamma-rays: stars – gamma-rays: diffuse background – methods: data analysis – methods: statistical

1 INTRODUCTION
Identifying astrophysical γ-ray sources is one of the main goals of the Fermi Gamma-ray Space Telescope. While the updated 10-year source catalogue, 4FGL-DR2 (the 4FGL-DR2, Ajello et al. 2020a), hereafter the 4FGL, identifies 279 sources as pulsars, and the 12-year catalogue, 4FGL-DR3, identifies 30 more γ-ray pulsars (Abdollahi et al. 2022), this is less than 10% of the pulsars in the Australia National Telescope Facility (ATNF) pulsar catalogue (Manchester et al. 2005).

With over fifteen years of almost continual all-sky survey data, the Fermi Large Area Telescope (LAT) is now becoming sensitive to pulsars that are below the putative γ-ray pulsar “death line” in spin-down luminosity E \(\lesssim\) 10^{33} erg s\(^{-1}\) (Smith et al. 2019) where the majority of rotationally-powered pulsars reside. Studies of faint γ-ray pulsars (e.g., Hou et al. 2014; Bruel 2019) illustrate both the potential and the difficulties of pushing the limits of sensitivity. Bright backgrounds, steep source spectra with low energy cutoffs, and broad, hard-to-discriminate pulsars make identifying low-luminosity pulsars challenging. Nevertheless, it is worth trying to detect this population because they are found in a corner of pulsar parameter space predicted to probe rare alignment geometries and potentially novel emission mechanisms (Smith et al. 2019).

Stacking methods in gamma-ray observations date back to the operational era of the Compton Gamma Ray Observatory’s Energetic Gamma Ray Experiment Telescope (EGRET) when stacking studies on galaxy clusters (Reimer et al. 2003) and galaxies (Cillis et al. 2004, 2005, 2007) were carried out. Other gamma-ray observatories like the High Altitude Water Cherenkov (HAWC) gamma-ray observatory used a similar joint-likelihood analysis (Albert et al. 2021) that resembles the stacking techniques developed for EGRET and Fermi-LAT. A significant amount of γ-ray flux from the known pulsars not yet detected in γ-rays should be present in the Fermi-LAT data. Authors such as Huber et al. (2012) proposed it should be discoverable using stacking techniques, but such techniques have up until now been unsuccessful in probing this population (McCann 2015). In contrast, stacking techniques that aggregate photon counts have been successfully employed in surveys of galaxy clusters (Dutson et al. 2013; Griffin et al. 2014; Prokhorov & Churazov 2014; Reiss & Keshet 2018), providing strong upper limits or detections of galaxy clusters as a population. More recently, an improved likelihood-stacking technique was developed and successfully characterized the faint blazar and starburst galaxy populations at GeV energies (Ajello et al. 2020b; Paliya et al. 2019), pushing well beyond the point-source detection limit of the LAT. This study extends these kinds of likelihood-stacking methods to the pulsar population with a new approach that stacks the measured Test Statistic (TS) values for pulsar fields, and, in so doing, provides some of the first measurements of this undetected population. This new method quantifies the significance of a stack observation and, when applied while varying spectral parameters, characterizes the γ-ray emission of the pulsar population under investigation. Such information can be used...
to inform how undetected populations of pulsars contribute to the diffuse γ-ray background, the γ-ray excess observed at the Galactic center, and γ-ray ﬂux from globular clusters.

With this project, we show evidence for the existence of a population of sub-threshold γ-ray pulsars and begin the investigation of this elusive population. We probe pulsars with high duty cycles with greatly improved sensitivity of the stacking methods. Furthermore, without assuming any beaming geometry of the sample pulsars, this work can potentially characterize un-pulsed γ-ray emission that may be produced by this population. The organization of this paper is as follows: In § 2, we describe our list of targets, the observational methods, and the stacking techniques. We outline the results of our analysis in § 3, where we summarize the new candidate population of sub-threshold pulsars including information that can be extracted from our stacking analysis. We also analyze control ﬁelds to validate the results. The overall properties of these sources are placed in context with the known pulsar population in § 4. This is followed by a brief conclusion and ﬁve short appendices which detail justiﬁcations for the speciﬁc choices in the methodology detailed in the main text and additional analysis of particularly signiﬁcant candidate sources.

2 OBSERVATIONS AND METHODS

We identify a sample of undetected pulsars in relatively uncrowded parts of the γ-ray sky to perform our stacking analysis. These locations are then analyzed using specialized software for Fermi LAT observations and the results are stacked as a sum of the Test Statistic (TS, deﬁned in § 2.2), which is derived from the likelihood of a source detection at a given location. Normal thresholds for detection (nominally TS > 25 as seen in the 4FGL) do not apply in such circumstances as, on average, the Fermi data tend to yield slightly positive TS values indicating that any random point in the sky is somewhat favored to be modeled as having a source. To better understand the signiﬁcance of a TS stack, a comparison to control ﬁelds—locations on the sky randomly selected to be similarly distributed as our test sample—is provided.

2.1 Data Selection

We choose target pulsars from the ATNF pulsar catalogue version 1.64, November 2020¹ that were not in the 4FGL catalogue using the psr py Python package² (Pitkin 2018). To optimize sensitivity, we avoid pulsars located in regions with complicated γ-ray emission backgrounds, such as the Galactic plane and bulge, and regions near extremely bright γ-ray sources. We therefore limit the search to high latitudes (|b| > 20°). Detailed reasoning for selecting this latitude cut is described in Appendix B. We also avoid all globular clusters and the Magellanic Clouds. A total of 362 pulsar positions met these initial selection criteria³.

Our target ATNF pulsars—along with the full ATNF catalogue and 4FGL pulsars—are plotted in Fig. 1. This high latitude study, while aimed primarily at reducing the background systematics, also probes a unique region of parameter space populated by older, transitional, and recycled millisecond pulsars (MSPs). The surveyed source distribution also more closely reﬂects the overall distribution of objects in the ATNF compared to those in the 4FGL.

Over twelve years of data from the LAT were used in this survey. The third revision of the PASS8 data, P8R3, released on Nov 26, 2018 was used along with the 4FGL and diffuse background models (Abdo et al. 2009). Data analysis in this work is performed with Fermipy (Wood et al. 2017) version 1.0.1⁴ based on the Fermi Science tools Anaconda distribution (Anaconda 2020) version 2.0.8⁵. Data were chosen between Mission Elapsed Time (MET) 239560000s (MJD = 54682) and 641950000s (MJD = 59340). Each region of interest (ROI) is a 21.2° × 21.2° square, which corresponds to a ~15° radius ROI. The data were also ﬁltered using a zenith angle cut of 90° to avoid bright emission from the Earth. Good time intervals were chosen with conditions DATA_QUAL==1 && LAT_CONFIG==1. The data were binned in 30 logarithmically-spaced energy bins (unbinned analyses for such a large duration of time are computationally prohibitive). An all-sky livetime cube and all-sky exposure cube⁶ within the speciﬁed time range from above were created and utilized for all ROIs.

Some previous studies have uncovered potential systematic effects studying faint populations at low energies (Paliya et al. 2020; Principe et al. 2021). We took a data-based approach to quantify this effect by conducting the likelihood analysis in two different energy ranges. Our primary stacking analysis only uses photons in the energy range of 300 MeV to 100 GeV to constrain the spectral energy distributions (SEDs) and calculate ﬁnal ﬂuxes. An analysis that includes additional photon data from 100 MeV to 300 MeV is found in Appendix D.

We also note that raising this lower energy limit further, e.g., to 1 GeV, would improve the differentiation of target sources and the background even more, but at too great a cost. The average cutoff energy of all Fermi detected pulsars is 830 MeV, so such a restriction eliminates most of the γ-ray signal we are attempting to uncover.

We also have created a library of 540 control ﬁelds from which we pull samples to validate the likelihood results, identify any systematic effects, and estimate the background signal. The control ﬁelds have the same radius of 15° and are drawn from locations within the same Galactic latitude and longitude distributions as the target pulsars. The center of each control ﬁeld is chosen to be at least 1° away from any known 4FGL source to be more certain that any measured signal can be attributed solely to ﬂuctuations in the background.

2.2 Fermi-LAT Data Analysis

Fermipy automatically reads the currently available version of the 4FGL catalogue and creates a model ﬁle for each ROI, along with the Galactic diffuse background emission and the isotropic background. We model an additional source at the center of each ROI as a power law with an exponential cutoff (PLEC) as the standard γ-ray spectral shape for pulsars (Abdo et al. 2013). For the model parameters, prefactor, index1, and cutoff are set free to be ﬁtted, while scale is set to 1000 MeV and index2 is set to 1.0 throughout the analysis. We set free the spectral parameters of all sources within 7.5° from the ROI center, and the diffuse background models. The pulsar positions are ﬁxed at the values reported in ATNF catalog. We note that some of the pulsars in the sample do not have a timing

¹ https://www.atnf.csiro.au/research/pulsar/psrcat/
² https://psrpy.readthedocs.io/en/latest/
³ the complete list of pulsars in this study can be found at: https://github.com/yuzhesong/stacking_funcs_tutorial/blob/main/pulsars_results_2023.csv
⁴ https://fermipy.readthedocs.io/en/latest/
⁵ https://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/binned_likelihood_tutorial.html
⁶ https://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/binned_likelihood_tutorial.html
Stacking of Gamma-ray Pulsars

Figure 1. $P$-$\dot{P}$ diagram of the pulsars surveyed: sub-threshold sources that were stacked (green diamonds), Fermi catalogue pulsars (blue dots) and ATNF pulsars (gray dots). Lines of constant characteristic age and $\dot{E}$ are indicated. The right side histogram shows the distribution of $\dot{P}$ for all the pulsars for which this has been measured; the top panel histogram shows the distribution of $P$ for all the pulsars including those not appearing on the scatterplot due to a lack of $P$ measurements. The histograms have the same color coding as the symbols, and the color bars are placed on top of each category. Out of the entire pulsar population, 482 pulsars do not have $P$ measurements and 26 of these do not have period measurements.

Position and may have relatively uncertain coordinates. However, it is beyond the scope of this work to provide precise locations of these pulsars, hence we fix the coordinates of the pulsars throughout our analysis. This process yields the TS value of each pulsar, defined as $TS = 2 \log \frac{L}{L_0}$, a comparison of the log likelihoods of the source model and the null hypothesis (no source, $L_0$). Usually, $\sqrt{TS}$ is used as a rough estimate of the source significance.

Additional unmodeled sources with $TS > 25$ were identified using the `find_sources` function in Fermipy, assuming they are power law sources.

2.3 Stacking Methods

In $\gamma$-ray observations, counts maps of each ROI can be overlaid and added up to create a stacked counts map. Broadly speaking, the resulting “stack” can be processed through the likelihood analysis to estimate the signal significance. Stacking photon counts has been applied to varied populations such as galaxies, galaxy clusters, and flare stars (Cillis et al. 2004, 2005, 2007; Griffin et al. 2014; Song & Paglione 2020). Huber et al. (2012) was among the early efforts to stack Fermi-LAT data by comparing stacked signal and background counts to calculate their likelihoods and the TS value of any detection. Direct likelihood stacking has also been conducted on blazars (Paliya et al. 2019, 2020), star-forming galaxies (Ajello et al. 2020b), and young radio galaxies and quasars (Principe et al. 2021). These studies not only obtained sensitive upper limits on the $\gamma$-ray emission, but have also made robust detections of some populations, plus a characterization of their optimal average spectral parameters. This study takes these methods and extends them to the undetected pulsar population described in § 2.1 while adopting a new interpretation of the cumulative TS stacks for characterizing the significance of a population’s potential signal over background and the TS maps of sources in spectral parameter space.

To test the existence of an undetected population, the first method used in this study is to consider simply adding up the TS values...
of all the sources of interest (SOIs) and comparing that sum to the cumulative TS of a population of control field test sources. Here we assume that a TS stack of SOIs, when compared to a TS stack of ROIs of blank sky positions, will result in a higher cumulative TS value if there is an underlying population of γ-ray sources emitting below the detection threshold. Stacking TS values is a much more sensitive measure than simple counts stacks, which are overwhelmed by noise given the low counts. The calculation proceeds as follows: For \( N \) sources, each with a TS value \( TS_k \), the cumulative TS value is

\[
TS = \sum_{k=1}^{N} TS_k = 2 \times \sum_{k=1}^{N} \ln L_{k,SOI} - 2 \times \sum_{k=1}^{N} \ln L_{k,null} \tag{1}
\]

which is a comparison between the collective maximum source likelihood and the collective null likelihood. Similarly, the cumulative TS value for \( N \) control field test sources is

\[
TS_{cf} = \sum_{k=1}^{N} TS_{k,cf} = 2 \times \sum_{k=1}^{N} \ln L_{k,cf} - 2 \times \sum_{k=1}^{N} \ln L_{k,null} \tag{2}
\]

Subtracting Eq. 2 from Eq. 1 effectively cancels the summed null term given appropriately chosen control fields. Thus, we can quantify the significance of the stack by introducing \( \Delta(TS) \)

\[
\Delta(TS) = 2 \times \sum_{k=1}^{N} \ln L_{k,SOI} - 2 \times \sum_{k=1}^{N} \ln L_{k,cf} \tag{3}
\]

If the stacked SOI TS values are consistently higher than the control field test source TS values, we may conclude that a signal exists, and the \( \Delta(TS) \) can be used to estimate the significance of the population, considering the control field likelihood term as the null.

One complication to address is that there is considerable variation in the distribution of TS values at various locations. Thus, a naive \( \chi^2 \) interpretation of the \( \Delta(TS) \) value may overstate the significance of a stack’s signal (see the left panel of Fig. 2 for the distribution of TS values for SOIs and control field ROIs and how they differ from the \( \chi^2 \) distribution). To account for this potentiality, a bootstrap resampling scheme is adopted to provide a conservative estimate of the uncertainties. Given \( N \) sources in either population in the stack, we randomly resample a given population \( N \) times, and calculate their cumulative TS distributions as a function of the stacked number of ROIs. This bootstrapping procedure is repeated \( M \) times, and the average and standard deviations of the cumulative TS values are calculated.

The entire process described above—a random sampling of \( N \) sources \( M \) times—is repeated for a total of \( P \) realizations. This process returns the distribution of the mean final cumulative (stacked) TS values for the population. Multiple sampling of a population that exceeds \( N \) (as in the case of our control fields) also yields a measure of the distribution for the final point of the stack. If a stack of sources is detected, the distributions of the cumulative TS value of the control fields should be statistically distinguishable from the final cumulative TS value of the source stack.

In the right panel of Fig. 2, we quantify the significance of this signal by measuring the separation of the two stacks. For the target pulsars, the uncertainty of the stack is estimated by stacking randomly re-ordered ROIs which creates a characteristic spindle-shaped envelope (shown in blue) as the variance increases halfway through the stack and then decreases until the final cumulative TS value. In contrast, for control fields (shown in green), the bootstrap resampling method described above yields an uncertainty that increases with the number in the stack. The \( \Delta(TS) \) between the final stacks can be used to roughly estimate the separation between the two populations, and \( \sqrt{\Delta(TS)} \) between the bootstrap uncertainty of the control fields can be used to estimate the significance of the separation.

As a second method, we investigate the spectral properties of this pulsar population by performing a TS stack in spectral parameter space similar to the likelihood stack in Paliya et al. (2019) and other works mentioned above. A TS map of the pulsar in spectral parameter space is created as described below. A point source with a PLEC spectrum is placed at the location of each pulsar fixing flux, spectral index, and cutoff energy. Only the background isotropic and Galactic diffuse normalizations are free to be fit by Fermipy’s fit function. This process returns a log-likelihood for the ROI given these spectral parameters for the central source. We repeat this process over a grid of flux and index values, keeping the cutoff energy fixed at 823.3 MeV to lower the degrees of freedom to 2. This cutoff energy is the median of all pulsars recorded in 4FGL-DR2. A detailed justification for fixing the cutoff energy is given in Appendix E. TS maps in flux-index parameter space for each ROI are made by subtracting the log likelihood of the grid point representing the null, which is chosen to be at the lowest flux \( (7.2 \times 10^{-12} \text{ ph cm}^{-2} \text{ s}^{-1}) \) and spectral index of \( -4 \), then multiplying this result by 2. The stacked TS map is then made by summing the individual TS maps. An over abundant amount of control fields is utilized for this study. To obtain the control field stack, we randomly choose the same number of control fields as the pulsar ROIs and add up all their TS maps, repeating this 1000 times, and averaging. Validation of this method is done by performing the analysis on a pulsar from the sample list, as described in § A.

3 RESULTS

3.1 Cumulative TS Stack

Fig. 2 shows the TS distributions of the central sources in the target ROIs measured above 300 MeV. The likelihood analyses for 9 of the pulsars in our sample did not converge and were omitted here. For comparison, the TS values for test sources at the centers of the control fields are also shown. The \( \chi^2/2 \) distribution for three degrees of freedom is also shown, which should be equivalent to the theoretical null (Mattox et al. 1996). Some of the pulsars have TS values above 25 and are further discussed for the possibility of detection in Appendix C.

To evaluate the uncertainty of the TS stacks, we did a bootstrap resampling of all ROIs \( M = 100 \) times for a total of \( P = 1000 \) realizations. The average cumulative TS value of the control fields is shown in two panels on the right of Fig. 2 along with their standard deviations. The cumulative TS value for the pulsars, with a standard deviation estimated from the bootstrap resampling technique are shown in the same panels. There is a statistically significant difference between the cumulative TS values of the pulsars and the control field test sources, \( \Delta TS > 200 \).

We quantify the significance of the stacked signal in a few different ways. Firstly and most intuitively, a Kolmogorov-Smirnov test (Kolmogorov 1933; Smirnov 1948) comparing ten thousand bootstrap resamplings of each distribution results in a statistic of \( D = 0.67 \), verifying that the two distributions are distinct with near statistical certainty. Secondly, we compare the final cumulative TS values of the pulsars and control field test sources. For the control fields, the final cumulative TS value is around 940 following the bootstrapping procedures. The mean bootstrap uncertainty is 82 and is represented by the green shaded area in Fig. 2. For the pulsars, the final cumulative TS value is 1144. The mean bootstrap uncertainty is 108 and is represented with the blue shaded area in Fig. 2. Between the
Figure 2. Results of TS stacking. Top left: TS values of 362 target pulsars from analysis between 300 MeV and 100 GeV. Bottom left: TS values of 362 randomly sampled control field test sources. The histogram values and the error bars are the averages and standard deviations after 100 samplings of the 540 control fields. The orange histograms represent the $\chi^2/2$ distribution for 2 degrees of freedom. The red dashed vertical line represents TS = 25, the usual detection threshold. Top right: Cumulative TS values of the target pulsars (blue) and control field test sources (green). The light blue shaded area represents randomly reordered stacks of the pulsar TS values. The light green shaded areas represent uncertainties of the stack for control fields estimated from bootstrap resampling of the stacks. Bottom right: The final mean cumulative TS value for pulsars (vertical blue line) and control fields (green histogram). The dashed green line represents the mean of the distribution. The dotted lines represent the mean of the standard deviations of the 1000 realizations of the bootstrap resampling process.

3.2 Parameter Space Stacking Results

The results of the parameter space stacking are shown in Fig. 3. Each panel indicates the location of a peak TS value, and the contours are within $\Delta TS = 13.2, 20.7$, and 30.3 of the peak, which correspond to idealized 3, 4, and 5$\sigma$ likelihoods. The left panel shows the result of all 353 pulsar ROIs stacked, and a similar analysis of control field test sources is shown for comparison.

For the pulsar ROIs, the maximum TS value of 300 occurs for a flux of $1.70^{+1.41}_{-0.68} \times 10^{-10}$ ph cm$^{-2}$ s$^{-1}$ and a spectral index of $-0.56^{+0.56}_{-0.93}$. Using 5$\sigma$ uncertainties. The sensitivity of the LAT for these latitude ranges is $3-5\times10^{-10}$ ph cm$^{-2}$ s$^{-1}$, about 2-3 times higher than the fluxes we are able to probe here. The spectral index is typical for 4FGL pulsars (Smith et al. 2019). For the control fields, the maximum TS value of 349 occurs at a flux of $3.94^{+2.04}_{-1.22} \times 10^{-10}$ ph cm$^{-2}$ s$^{-1}$ and a spectral index of $-2.63^{+1.12}_{-1.38}$. At the location in parameter space where the peak TS value for the pulsars occurs, the $\Delta(TS)=28$ between the pulsars and control fields, which corresponds naively to 4.8$\sigma$ for 2 degrees of freedom. We note that although a peak of similar TS and flux occurs in the parameter space stack for the control fields, that peak corresponds to a significantly softer spectral index than for a typical pulsar. As such, the diffuse background and/or unmodeled sources in Fermi data can be systematic effects or signal contamination in stacks that resemble the stack of the control fields. Very soft spectra generated from stacking should be treated with suspicion as our analysis and previous work (Paliya et al. 2020) show this may be generated when stacking blank sky.

3.3 Stacked Spectral Energy Distributions

We stack the signals in different photon energy bins to generate a stacked spectral energy distribution (SED) for pulsars with properly calibrated distances (Fig. 4). To assess the SED of the collection of pulsars properly, the pulsar fluxes are recalibrated to the same distance of 1.695 kpc (the average of the distances in the sample). Here we bin the data into three energy bins: 300 MeV to 1 GeV, 1 GeV to 10 GeV, and 10 GeV to 100 GeV. While pulsars are known for their distinctively curved SEDs, in each energy bin a power law
Figure 3. Parameter space stacks of pulsars and control fields with various weights. Left: parameter space stack of 353 ROIs with candidate pulsars. The white X marks the optimal parameters with the largest TS value of 300. Right: Parameter space stack for control fields test sources, produced by randomly choosing 353 ROIs out of 540 control fields and repeating this random selection 1000 times, then averaging over the randomisations. The white X marks the optimal parameters with largest TS value of 349. The brown, red, and white contours indicate 3σ, 4σ, 5σ from the maximum in all panels.

(PL) fit is used, as is standard practice\footnote{https://fermipy.readthedocs.io/en/latest/advanced/sed.html}, to simplify the analysis and reduce the degrees of freedom. The binned parameter space stack converges for 218 of the pulsar fields. We first obtain the best-fit PL index for the sample pulsars from the parameter stack. Fixing this PL index, the flux in each bin is then calculated from the best fit flux value of the largest TS in the stacked parameter space map, and uncertainties are evaluated from the 3σ spread of the flux value from the peak. The TS map of the highest energy bin does not converge in parameter space, so an upper limit is given. The stacked SED further corroborates the pulsar-like properties of the sub-threshold sources. To examine the effect of background contamination on the SED, the stack of pulsars without distance calibration is compared to that of the control fields. These results have similar parameter stack structures in all energy bins, although the maximum cumulative TS values vary. In particular, between 1 and 10 GeV, the parameter stack of the pulsars has a much higher peak TS than the control fields.

To show that this stack of pulsar ROIs is consistent with previous pulsar observations, we compare the stacked SED to the previous pulsar stacking results of McCann (2015). In Fig. 4, the green dashed-dotted line is their average best-fit spectrum for young pulsars in the catalogue with a PL index of −1.46 and a cutoff energy at 3.6 GeV; the green dotted curve is the average best-fit spectrum for the catalogue MSPs, with a PL index of −1.59 and a cutoff energy at 3.54 GeV. Our stacked SED is about an order of magnitude less luminous than the average MSP and two orders of magnitude less luminous than the average young pulsar, which shows the sensitivity of the stacking technique. To compare our results to the average SEDs of young pulsars/MSPs, the previously modeled SEDs are plotted on a different scale to match the positions of the data points. This result is consistent with the pulsars in our stacking survey being less luminous and with generally lower $\tilde{E}$, which is explored further in § 4.2.

4 DISCUSSION

4.1 Rate of False Detection/Controlling for the Background

In this study, we find that five target ROIs and two control field test sources have TS values exceeding the nominal source-detection threshold of TS > 25 (Fig. 2, right panel). Modulo issues with interpreting the tails of the distribution of TS values, we can consider the control fields as effectively probing the false detection or background floor in this study.

We also examine simulated data to understand if the control fields with high TS could be caused by uncertainties in modeling the background and/or other sources. The simulated observations are generated by the simulate_roi function in Fermipy. The central source is left out of the model when creating the simulated observation, and the normalization of the isotropic background is adjusted to test the effects of varying counts noise. The simulated observations of each of the control field ROIs are then fed through the analysis pipeline with the central source added back to the model, and the TS value of the central source is measured. Without altering the background normalization, the resulting TS distribution is indistinguishable from the null, $\chi^2/2$, which is inconsistent with the control field results in Fig. 2 and indicates that systematic observable effects are causing a high control-field signal. TS mapping and stacking is thus highly sensitive to background modeling. An increase of only 0.1% in the isotropic background normalization yields test source TS values that
and used the default ATNF distance modeled with YMW2016 (Yao et al. 2017) with a 30\% uncertainty to estimate their luminosities and uncertainties. We consider a beaming factor \( f_{\text{beam}} = 1 \), where the luminosity is the energy flux multiplied by \( 4\pi d^2 f_{\text{beam}} \). We estimated the luminosity of candidate pulsars based on the results from the parameter space stacking in \S 3.1. The flux and photon index are obtained from the stacked TS maps in each \( E \) range and integrated to get the energy flux in the 100 MeV to 100 GeV energy range. Uncertainties in energy flux are estimated from the 3\sigma spread of the parameters. Luminosity is then calculated by scaling to the mean of distance squared of the sub-threshold sources. The 7 pulsars without a YMW2016 distance or the 51 pulsars set to the flagged value of 25 kpc based on the dispersion measure are again excluded.

The luminosity and \( E \) of all 4FGL-DR3 pulsars and the candidate pulsars are plotted in Fig. 6. The sample pulsars appear to effectively extend the luminosity-\( E \) dependence to low \( E \). And provide more evidence for the suggestion that at low \( E \) the heuristic relation of \( L_\gamma \propto \sqrt{E} \) could change to an even stronger dependence and a very high efficiency (Smith et al. 2019).

There are two issues that might impact the results of this analysis. Firstly, pulsar distances in the YWM2016 survey are quite uncertain. These distances, when compared to some existing parallax measurements in the ATNF catalogue, can be a factor of few different, resulting in as large as an order of magnitude difference in the final luminosity estimation. Secondly, the \( P \) measurements of MSPs in the ATNF catalogue are not corrected to the Shklovskii effect (Shklovskii, I. S. 1970). Subsequently, the estimated \( E \) of MSPs are not accurate. However, in this work, only 22 out of the sample pulsars in the stacked sample are MSPs. Especially when considering the fact that none but five of the stacked pulsars are individually detected, and none of these five pulsars are MSP, corrections for this effect are not crucial in this work.

Many models of \( \gamma \)-ray emission from pulsars indicate that the emission mechanism leading to the pulsed signal may turn off at low \( E \) (Watters et al. 2009). Low \( E \) also precludes them from being pulsar wind nebula candidates. This survey was designed specifically to have clean backgrounds so the \( \gamma \)-ray source should originate with the neutron star itself, similar perhaps to the presumed magnetospheric off-peak emission seen in some pulsars (Abdo et al. 2013). It is worth noticing that the majority of \( \gamma \)-ray pulsars are not detected with off-peak emission. This work provides better sensitivity that would likely to detect said weak off-peak emission. We investigate a simple scenario where the isotropic \( \gamma \)-ray emission is from an inverse Compton halo located near the light cylinder. Using the naima code (Zabalza 2015; Khangulyan et al. 2014), we assume a seed photon field from photospheric thermal emission at \( 10^6 \) K. Integrating over all scattering angles, inverse Compton scattering from an electron distribution with a power law index of \( -1.8 \) and cutoff energy of 4 GeV generates a spectrum that is consistent with the fluxes and SEDs of the detected and sub-threshold sources, as shown in Fig. 4. The required normalization depends on the product of the non-thermal electron density and photon density (which depends on the characteristic scattering distance from the stellar photosphere). If this distance is comparable to the light cylinder radius, \( \gtrsim 10 \) neutron star radii, then the lepton densities required to create such a spectrum are orders of magnitude below the Goldreich-Julian density (Goldreich & Julian 1969). Recent development of pulsar magnetospheric models, such as the current sheet model (Kalapotharakos et al. 2018), indicate the existence of such leptons at the equatorial plane of a pulsar around the light cylinder, which can provide the population for the IC scattering process.

4.2 Luminosities

Unsurprisingly, we find that, as a population, the undetected pulsars in our stack have significantly lower luminosities than the 4FGL-detected pulsars. This study allows us to probe the low rotation-power \( (E) \) regime. Most of the pulsars in our sample also do not have \( \gamma \)-ray pulsations detected. Comparing the implied luminosities of our stack to the known \( \gamma \)-ray luminosities allows us to extend the known correlation between \( \gamma \)-ray luminosities and \( E \) (Abdo et al. 2013).

During the course of this study, the Fermi-LAT collaboration updated the point source catalogue to 4FGL-DR3 containing \( \sim 30 \) more pulsars. For a fairer and more complete comparison, we use all the pulsars in DR3. We extracted the energy flux from 4FGL-DR3...
In this work, we make no assumptions on beaming or emission geometry of the pulsars. Out of the 362 pulsars in our sample, 105 have either no $\dot{P}$ or no $P$ measurement, meaning that we cannot readily search for $\gamma$-ray pulsations from them, and we cannot assume any emission geometry. The fraction of pulsars with $\gamma$-ray pulsations is a few percent to 2/3 for different $\dot{E}$ ranges (Abdo et al. 2013). Detectability depends not only on the pulsar luminosity and its distance from the Earth, but also the beaming geometry, e.g., the size of the beam and the viewing angle. How do we determine if the stacked signal is from faint $\gamma$-ray pulses beamed towards the Earth from a small subset of our sample, or some unknown isotropic $\gamma$-ray emission emitted by all pulsars?

We first examine the correlation between pulsar $\dot{E}$ and their TS values from the single ROI binned likelihood analysis. As shown in Fig. 7, there is no obvious correlation between TS and pulsar $\dot{E}$. Since pulsed $\gamma$-ray luminosity is generally dependent on $\dot{E}$, the stacked signal is not dominated by just a few high $\dot{E}$ pulsars. Instead these pulsars may have weak, unpulsed or very broadly pulsed gamma-ray emission that originates, for example, from the equatorial region around the light cylinder of the pulsar, or is more isotropic.

Similarly, there is no significant difference in the parameter space stacking results after grouping by $\dot{E}$ (Fig. 8). The peak TS values in each $\dot{E}$ bin are very similar, and their parameter uncertainties overlap considerably. This result reinforces that the $\gamma$-ray emission mechanism from the three distinctive pulsar populations might be very similar. Given the large error bars on the luminosity estimation of these populations, we cannot say with any further certainty that this mechanism is $\dot{E}$ dependent or not. We defer the exploration of low $\dot{E}$ pulsar $\gamma$-ray emission mechanism to future study.

Previous searches for $\gamma$-ray pulsations from pulsars (e.g., Smith et al. 2019) did not detect pulsations from a large fraction of high $\dot{E}$ pulsars. This result might be dominated by the emission geometry rather than their $\gamma$-ray luminosities. From our sample of pulsars, 174 were phase-folded in Smith et al. (2019). Here, we try to correlate our results with the $H_{\text{max}}$ values introduced in Smith et al. (2019). $H_{\text{max}}$ is defined as the maximum H-test value when H-test is performed in six different ways (refer to Sec. 3 in Smith et al. (2019) for further details). The highest $H_{\text{max}}$ was 1212 for J0514-4408 which is a 4FGL-DR3 recorded pulsar, an indication of a significantly detected periodic signal. We note that 5 pulsars have $H_{\text{max}} > 25$, including J0922+0638, 148 have $H_{\text{max}} < 25$, and 21 did not return a valid $H_{\text{max}}$. Again there is no significant difference between the parameter stacks of the subgroups of pulsars that were phase-folded or not, or resulted in high or low $H_{\text{max}}$ values. The stack of the 5 pulsars with $H_{\text{max}} > 25$ seems to exhibit a significant peak with TS = 30 as shown in the bottom left panel of Fig. 8. It is worth noting that the parameter space stack of two of these five pulsars did not converge, and the parameter space stack of J0922+0638 as shown in Fig. 8.1A dominates this stack. The stack for the 148 low $H_{\text{max}}$ pulsars shows a peak with TS = 120. Randomly selecting 5 pulsars out of this list indicates a small chance (< 2%) of a peak TS larger than 30. While the parameters are not well constrained in either case, this result seems to imply a correlation between the peak TS and the $H_{\text{max}}$ values of the pulsars being stacked, but with a very small number of high $H_{\text{max}}$ pulsars, the correlation is not significant. One can also search for the existence of weak, isotropic $\gamma$-ray emission from pulsars by examining the off-peak phases of detected pulsars, which is reserved for future study.

5 CONCLUSIONS

In this work we examined the $\gamma$-ray emission from 362 locations that coincide with ATNF pulsars using Fermi-LAT data. Our sample ranges widely from young to old pulsars, magnetars, MSPs, and low $\dot{E}$ pulsars.
ACKNOWLEDGEMENTS

The authors would like to express gratitude towards the reviewer, D. Smith, for his constructive feedback and comments on this work. This work was supported in part by NSF grants AST-1831412, AST-2219090, and AST-1852360, as well as PSC-CUNY award #63785-00 51, and the CUNY Research Scholars Program. The corresponding author is also supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project number CE170100004. The authors thank D. Zurek, M. Bailes and S. Stevenson for very useful discussions. This project made use of computational systems and network services at the American Museum of Natural History supported by the National Science Foundation via Campus Cyberinfrastructure Grant Awards # 1827153 (CC* Networking Infrastructure: High Performance Research Data Infrastructure at the American Museum of Natural History).

DATA AVAILABILITY

Fermi-LAT photon data are available through Fermi-LAT data server. Fermi-LAT analysis results (the output .xml or .npy files from Fermi-ny) of target pulsars and control fields can be shared upon request to the first author. Post-processing analysis scripts will be shared upon request to the first author, given that a proper citation to this paper is provided in the work from the people that make the request.

REFERENCES

Abazajian, K. N. 2011, J. Cosmology Astropart. Phys., 2011, 010
Abdo, A. A., Ackermann, M., Ajello, M., Atwood, W. B., Axelsson, M., et al. 2009, ApJ, 703, 1249
Abdo, A. A., Ajello, M., Allafort, A., Baldini, L, Ballet, J, et al. 2013, ApJS, 208, 17
Abdollahi, S., Acero, F., Baldini, L, Ballet, J, Bastieri, D., et al. 2022, ApJS, 260, 2, 53
Albert, A., Alfaro, R., Alvarez, C., et al. 2021, ApJ, 911, L27. doi:10.3847/2041-8213/abf4dc
Ajello, M., Angioni, R., Axelsson, M., Ballet, J, Barbieri, D. et al. 2020a, ApJ, 892, 105A
Ajello, M., Di Mauro, M., Paliya, V. S., & Garrison, S. 2020b, ApJ, 894, 88
Albert, A., Alfaro, R., Alvarez, C., et al. 2021, ApJ, 911, L27
Aloisi, R. I., Cruz, A., Daniels, L., Meyers, N., Rookie, R. et al., 2019, ApJ, 875, 19
Anaconda Software Distribution, Anaconda Documentation, Vers. 2-2.4.0
Brue, P. 2019, A&A, 622, A108
Burke-Spolaor, S. & Bailes, M., 2010, MNRAS,402, 2, 855-866
Cillis, A. N., Hartman, R. C., and Bertsch, D. L., 2004, ApJ, 601, 142
Cillis, A. N., Torres, D. F., and Reimer, O., 2005, ApJ, 621, 139
Cillis, A. N., Reimer, O, and Torres, D. F, 2007, Ap&SS, 309, 51
Dutson, K. L., White, R. J., Edge, A. C., Hinton, J. A., & Hogan, M. T, 2013, MNRAS, 429, 2069
The Fermi-LAT Collaboration. 2017, arXiv e-prints, arXiv:1705.00009.
Griffin, R. D., Dai, X., & Kochanek, C. S. 2014, ApJ, 795, L21.
Goldreich, P. & Julian, W. H. 1969, ApJ, 157, 869. doi:10.1086/150119
Hobbs, G., Lyne, A. G., Kramer, M., Martin, C. E. & Jordan, C., 2004, MNRAS, 353, 4, 1311-1344
Hou, X., Smith, D. A., Guillemot, L., Cheung, C. C., Cognard, I., et al. 2014, A&A, 570, A44
APPENDIX A: VALIDATING PARAMETER SPACE ANALYSIS METHOD

As a proof of concept, we performed the same parameter stack scheme on J0922+0638 because it was not recorded in 4FGL-DR2, but was recorded in 4FGL-DR3 as 4FGL J0922.3+0644. DR3 gives a PL model of the source, with TS = 29, a power law spectral index of $\gamma = 2.0 \pm 0.14$, and a flux from 1 to 100 GeV to of $(1.19 \pm 0.28) \times 10^{-10}$ ph cm$^{-2}$ s$^{-1}$, which can be converted to $(8.2 \pm 3.0) \times 10^{-10}$ ph cm$^{-2}$ s$^{-1}$ for photons between 0.3 to 100 GeV.

The parameter space analysis of the pulsar in the left panel of Fig. A1, yields a peak TS = 28 with a $\gamma$-ray flux (300 MeV - 100 GeV) of $(3.0 \pm 2.6) \times 10^{-10}$ ph cm$^{-2}$ s$^{-1}$, consistent with the catalogue value, as well as a power law index of $\gamma = 1.6 \pm 0.12$, and a fixed cutoff energy of 823 MeV. We plotted the SED of the pulsar using the results of the spectral parameter space TS map, using the parameters from the catalogue as described in the last paragraph, and using the results from a standard binned likelihood analysis as tabulated in Table C1 (Fig. A1). These results are all consistent with each other within the uncertainties.

APPENDIX B: DETERMINATION OF LATITUDE CUT

In the main text, only the 362 pulsars that are 20° away from the plane are used in the stacks, which was motivated by analysing the impact of different latitude cuts on the stack sensitivity. This appendix describes the determination of this latitude limit, which derived from an analysis all pulsars located 5° away from the Galactic plane not already detected in $\gamma$-rays. For sources within 20° of the Galactic plane, only those 30° away from the Galactic center were analyzed. A total of 726 pulsars satisfy these selection criteria.

Fig. B1 shows the final cumulative TS value for the pulsars and control field test sources as a function of the lower latitude cut. The error bars on both curves are the standard deviation of the final cumulative TS values using the bootstrap technique described in § 2.3.

Figure A1. Left: TS map of PSR J0922+0638 (4FGL J0922.3+0644) in parameter space, following the parameter stacking analysis scheme described in § 2.3. The white X marks the peak TS value of 28 in the chosen region in parameter space and the black and red contours represent the 3$\sigma$ and 5$\sigma$ away from the peak. Right: Comparison of the SED of J0922+0638 generated using the catalogue parameters (red) as described in text, results from the standard binned likelihood analysis as recorded in Table C1 (blue), and parameter stack results from the left panel (green). The shaded areas each represent the 95% uncertainty in the flux estimation.

Figure B1. Cumulative TS values and uncertainties of the pulsars (blue line with error bars) and the corresponding control field test sources (green dashed line with error bars) with the same spatial distribution, plotted against the minimum latitude selection.
APPENDIX C: DISCUSSION OF PULSAR CANDIDATES WITH TS > 25

As mentioned above, five pulsars from our sample have TS > 25, the usual detection threshold. To further assess the γ-ray emission from these locations, we found the peak TS value each of these pulsars using the localize routine in fermipy. We also reoptimized the ROI. Their optimized locations and uncertainties, original locations and uncertainties from the ATNF, and SEDs, are shown in Fig. C0. Their spectral properties obtained in this work are shown in Table C1. The same analysis was performed on the two control field test sources that have TS > 25 and the results are also shown in Fig. C0 and Table C1. One candidate, PSR J0922+0638, has now been subsequently reported in the 4FGL-DR3 as a newly detected γ-ray pulsar.

PSR J0038-2501 and PSR J1705-04, have well fitted parameters and pulsar-like SEDs. PSR J1610-17, has a poorly fit index, and its cutoff energy is extremely low, which overall results in a very soft SED. PSR J2336-01 has a relatively soft spectral index, a largely unconstrained cutoff energy, and its SED is slightly more consistent with a simple power law spectrum rather than a PLEC shape. The curvature significance of each source is quantified by the TS_{curv} value included in Table C1.

We also examined the individual map for these five pulsars. J0922, J0038 and J1705 all have relatively good individual maps in parameter space. The other two, J1610 and J2336 have upper-limit/background like maps in parameter space.

Overall, PSR J1705-04 is potentially a good candidate for follow-up study to see whether its γ-ray flux can be confirmed to be associated with the pulsar. PSR J0038-2501 has a TS value as high as 55 before localization and reoptimization, the TS value is 33 after localization and reoptimization. It is only 0.04° away from 4FGL J0038-2459, a γ-ray-detected flat spectrum radio quasar, making it hard to distinguish the signal from the two sources. PSR J2336-01 and PSR J1610-17 seem more likely to be spurious or due to flux from an unassociated source near the line of sight given their soft and/or low curvature SEDs. Additionally, PSR J2336-01 is within 1° of three FSRQs in the catalogue: 4FGL J2335.4-0128, 4FGL J2338.0-00230, and 4FGL J2333.4-0133. However, the likelihood analysis pipeline could separate these signals. While phase-folded pulse detection is needed to confirm the pulsar nature of these candidates, attempting to establish γ-ray timing solutions is reserved for future work.

Two of these five pulsars have $\dot{E}$ measurements. Their efficiencies, $L_{\gamma}/E = 53$ and 334, respectively, which are extraordinarily large. These extremely high over-unity efficiencies could be an indicator that these gamma-rays do not originate from the pulsars. However, like all the other sample pulsars, the distances of these two are largely uncertain, a factor which impacts the estimated luminosity greatly. Additionally, both of these pulsars have $\dot{E}$ much lower than the death-line value of $10^{33}$ erg/s. Given the limited understanding of pulsar behaviours in this regime, they also could possibly have $f_{\nu}$ values very different from unity due to different emission geometries.

It might be worth noting that, another pulsar in our list, J1455-3330, that have slightly higher, yet still undetected signals (10 <TS< 25), is also now present in the 4FGL-DR3 catalogue.

For completeness, we note that there are also two TS > 25 control field test sources found in our library. Performing a similar analysis on them resulted in poorly localized positions and uncertain spectral parameters. The fitting results did not converge for the control field labeled as CF970. CF8, while seemingly well-localized and with a pulsar-like SED, has a suspiciously low index uncertainty, which usually indicates a poor fit. Thus, we consider this control field detection to be spurious as well.

APPENDIX D: STACKING RESULTS FROM 100 MEV

Recent work has revealed that due to the worse resolution and contamination from the backgrounds, parameter stacking that includes photons below 300 MeV may introduce a spurious population of extremely soft sources (Paliya et al. 2020). Here, we test this claim by comparing the stacking analysis results using the 300 MeV to 100 GeV data to those using the 100 MeV to 100 GeV data following the same prescription outlined in § 2. In Fig. D1, we plot the TS_{100} (TS value of a source obtained using data between 100 MeV and 100 GeV) distribution and the cumulative TS_{100} distribution of both the pulsar ROIs and control field ROIs. While the stacked pulsar signal still persists when examining the cumulative TS_{100} distributions, the difference from the control stack is less significant and the standard deviation of the mean final cumulative TS value is larger.

We also investigated the 100 MeV to 100 GeV pulsar and control field ROIs using the parameter space stacking technique. Only 315 of the pulsar ROIs converged at all points in parameter space using this method. In Fig. D1, we show the parameter stacking results of these 315 pulsar and control field ROIs that converged between 100 MeV and 100 GeV. The control field stack is again performed by randomly choosing 315 fields out of the 540, repeating this 1000 times, and averaging. The parameter space stacking of pulsars shows no significant difference from the results using 300 MeV to 100 GeV data in Fig. 3. The control fields stacking results, still have a softer spectral index than the pulsars, though slightly harder than the stack results between 300 MeV and 100 GeV. The average flux is also higher, enough to breach the detection threshold of Fermi and contribute significant spurious signal.

Overall, stacking the control fields from 100 MeV does seem to capture the noise/background contamination in the low energy range. Hence, we recommend using the higher energy range, due to the following additional considerations: 1) In our analysis, we used a curved PLEC spectrum for our sources instead of a simple power law. A larger systematic effect could occur for power law sources. 2) Above 300 MeV, the improved resolution of Fermi-LAT can greatly aid in localizing faint signals and in making less ambiguous source associations. 3) We also saw a decrease in the control field background signal when excluding the lower energies.

APPENDIX E: THREE-DIMENSIONAL PARAMETER SPACE STACK

With three parameters (prefactor, index and cutoff energy) for PLEC fitting, a complete parameter-space stack could include, in addition to the flux and index, the cutoff energy. It is a feature of the known population of γ-ray pulsars that the cutoff energies lie within a very small range, which is 1.7 GeV for MSPs, and 0.9 GeV for radio-loud γ-ray pulsars (Abdollahi et al. 2022). It is therefore reasonable to set the cutoff energy to a fixed value in our methods.

To further test the dependence on cutoff energy, we calculated the same parameter stack using data from 300 MeV to 100 GeV at 4 different cutoff energies with a sample selection of 138 pulsars from the target list. The results for cutoff energies of 655 MeV, 823 MeV, 1034 MeV and 1300 MeV, are shown in Fig. E1. There is a slight softening of the spectrum with increasing cutoff energy, which is consistent with the trends for 4FGL pulsars and the analytical relation between the two quantities derivable directly from the PLEC model. However, the 5σ range is large compared with the trend, and the peak TS value does not heavily depend on the choice of the cutoff energy. We therefore chose to fix the cutoff energy at the median.
Figure C0. Spectral energy distributions and TS maps of target pulsars and control field test sources with TS > 25. Fluxes with errorbars are shown in each energy bin with TS > 4, and the rest are shown with 95% upper limits. We also plot the best-fit PLEC (black solid line) spectral models from the likelihood analysis over the full energy range along with the uncertainties (gray shaded area), with the parameters given in Table C1. The red + and ellipse represent the radio position and uncertainty of each pulsar; the white x and ellipse indicate the optimal $\gamma$-ray location and 95% uncertainty.

value of all the PLEC 4FGL pulsars for the analysis presented in § 3.1, which also conveniently reduces the degrees of freedom.

This paper has been typeset from a \TeX/La\TeX file prepared by the author.
### Table C1. Candidate Pulsar and Control Field Detections

|                | RA (Fermi) | Dec (Fermi) | RA (radio) | Dec (radio) | Photon Flux | Index | Cutoff Energy | TS | $T_{\text{SNR}}$ | Luminosity | $E$ | Dist (kpc) |
|----------------|------------|-------------|------------|-------------|-------------|-------|---------------|----|---------------|------------|-----|-----------|
| J0038-2501     | 00:38:37.51 | -25:01:52.2 | 00:38:37.17 | -25:01:52.2 | 0.68 ± 0.22 | 1.87  | 2.52 ± 1.33  | 330| n/a           | 0.94 ± 0.60 | 0.47 | 0.064     |
| J0922+0638$^a$| 09:22:35:44 | 06:38:45:3   | 09:22:35:43 | 06:38:45:3  | 1.96 ± 0.24 | 1.87  | 2.26 ± 0.57  | 383| 2.2           | 3.3 ± 0.31  | 1.1  | 6.8       |
| J1040-17       | 10:40:17:56 | -17:40:26:3  | 10:40:17:56 | -17:40:26:3 | 3.04 ± 0.38 | 2.00  | 2.20 ± 0.51  | 303| 3.0           | 3.0 ± 0.51  | 2.5  | 2.07      |
| J1205-04       | 12:05:04:17 | -04:12:31:54 | 12:05:04:17 | -04:12:31:54 | 2.03 ± 0.44 | 2.00  | 2.11 ± 0.46  | 334| 3.1           | 3.1 ± 0.46  | 2.5  | 3.37      |
| J2336-61       | 23:36:19:04 | -61:55:33:25 | 23:36:19:04 | -61:55:33:25 | 4.18 ± 0.28 | 2.00  | 3.41 ± 0.42  | 1073| 3.3 ± 0.57   | 8.29 ± 0.52 | 2.5  | 3.36      |
| CPS            | 11:18:26:87 | 58:58:68:12 | 11:18:26:87 | 58:58:68:12 | 0.09 ± 0.03 | 1.87  | 1.49 ± 0.062 | 345| 3.3 ± 0.46   | 1.19 ± 0.52 | 2.5  | 3.39      |
| CF970          | 16:05:33:10 | -16:13:50:24 | 16:05:33:10 | -16:13:50:24 | 2.46 ± 0.39 | 1.87  | 3.77 ± 1.51  | 420| 5.9 ± 0.57   | 5.91 ± 0.71 | 2.5  | 3.38      |

$^a$ Reported in 4FGL DR3

$^b$ Burke-Spolaor & Bailes (2010)

$^c$ Hobbs et al. (2004)

$^d$ Sanidas et al. (2019)

$^e$ Power law model did not converge, indicating that this model is strongly disfavored.
Figure D1. Top left: $\text{TS}_{100}$ values of target pulsars using data from 100 MeV to 100 GeV data. Bottom left: TS values of control field test sources from analysis with 100 MeV to 100 GeV data. The orange histogram represents the $\chi^2/2$ distribution for 2 degrees of freedom. The red dashed vertical line represents $\text{TS} = 25$, the presumed detection threshold. Right: Cumulative $\text{TS}_{100}$ values of the target pulsars (blue) and control field test sources (green). The light blue curves represent stacks after one randomly re-ordering of the TS distribution of the pulsars. The light green shaded areas represent uncertainties of the stack for control fields estimated from bootstrap resampling of the stacks. Bottom left: Parameter space stack of 315 pulsars. The white X marks the optimal parameters with the largest TS value of 397, and the contour shows the $5\sigma$ range. Bottom right: Parameter stack of 315 control fields.
Figure E1. Parameter space stacking of 138 randomly sampled pulsar ROIs with 4 different cutoff energies of 655 MeV, 823 MeV, 1034 MeV and 1300 MeV, from top left to top right to bottom left to bottom right, respectively.