A STATISTICAL APPROACH TO RECOGNIZING SOURCE CLASSES FOR UNASSOCIATED SOURCES IN THE FIRST *FERMI*-LAT CATALOG

M. Ackermann[1], M. Ajello[2], A. Allafort[2], E. Antolini[3,4], L. Baldini[5], J. Ballet[6], G. Barbiellini[7,8], D. Bastieri[9,10], R. Bellazzini[5], B. Berenji[2], R. D. Blandford[2], E. D. Bloom[2], E. Bonamente[3,4], A. W. Borgland[2], A. Bouvier[11], T. J. Brandt[12,13], J. Bregeon[5], M. Brigida[14,15], P. Bruel[16], R. Buehler[2], T. H. Burnett[17], S. Buson[9,10], G. A. Caliandro[18], R. A. Cameron[2], P. A. Caraveo[19], J. M. Casandjian[6], E. Cavazzuti[20], C. Cecchi[3,4], Ö. Çelik[21,22,23], E. Charles[2], A. Chekhtman[24,68], A. W. Chen[19], C. C. Cheung[25,68], J. Chiang[2], S. Ciprini[4,26], R. Claus[2], J. Cohen-Tanugi[27], J. Conrad[28,29,69], S. Cutini[20], A. de Angelis[30], M. E. DeCesar[21,31], A. De Luca[32], F. de Palma[14,15], C. D. Dermer[33], E. do Couto e Silva[2], P. S. Drell[2], A. Drlica-Wagner[2], R. Dubois[2], T. Enoto[2], C. Favuzzi[14,15], S. J. Fegan[16], E. C. Ferrara[21], W. B. Focke[2], P. Fortin[16], Y. Fukazawa[34], S. Funk[2], P. Fusco[14,15], F. Gargano[15], D. Gasparrini[20], N. Gehrels[21], S. Germani[3,4], N. Giglietto[14,15], F. Giordano[14,15], M. Giroletti[35], T. Glanzman[2], G. Godfrey[2], I. A. Grenier[6], M.-H. Grondin[36,37], J. E. Grove[33], L. Guillemot[38], S. Guiriec[39], M. Gustafsson[9], D. Hadasch[18], Y. Hanabata[34], A. K. Harding[21], M. Hayashida[2,40], E. Hays[21], S. E. Healey[2], A. B. Hill[41], D. Horan[16], X. Hou[42], G. Jóhannesson[43], A. S. Johnson[2], T. J. Johnson[25,68], T. Kamae[2], H. Katagiri[44], J. Kataoka[45], M. Kerr[2], J. Knödlseder[12,13], M. Kuss[5], J. Lande[2], L. Latronico[46], S.-H. Lee[47], M. Lemoine-Goumard[48,70], F. Longo[7,8], F. Loparco[14,15], B. Lott[48], M. N. Lovellette[33], P. Lubrano[3,4], G. M. Madejski[2], M. N. Mazziotta[15], J. E. McEnery[21,31], J. Mehault[27], P. F. Michelson[2], R. P. Mignani[49], W. Mitthumsiri[2], T. Mizuno[34], C. Monte[14,15], M. E. Monzani[2], A. Morselli[50], I. V. Moskalenko[2], S. Murgia[2], T. Nakamori[45], M. Naumann-Godo[6], P. L. Nolan[2,71], J. P. Norris[51], E. Nuss[27], T. Ohsugi[52], A. Okumura[2,53], N. Omodei[2], E. Orlando[2,54], J. F. Ormes[55], M. Ozaki[53], D. Paneque[2,56], J. H. Panetta[2], D. Parent[57,68], V. Pelassa[39], M. Pesce-Rollins[5], M. Pierbattista[6], F. Piron[27], G. Pivato[10], T. A. Porter[2], S. Rainò[14,15], R. Rando[9,10], P. S. Ray[33], M. Razzano[5,11], A. Reimer[2,58], O. Reimer[2,58], T. Reposeur[48], R. W. Romani[2], H. F.-W. Sadrozinski[11], D. Salvetti[19], P. M. Saz Parkinson[11], T. L. Schalk[11], C. Sgrò[5], M. S. Shaw[2], E. J. Siskind[59], P. D. Smith[60], G. Spandre[5], P. Spinelli[14,15], D. J. Suson[61], H. Takahashi[52], T. Tanaka[2], J. G. Thayer[2], J. B. Thayer[2], D. J. Thompson[21], L. Tibaldo[9,10], O. Tibolla[62], D. F. Torres[18,63], G. Tosti[3,4], A. Tramacere[2,64,65], E. Troja[21,72], T. L. Usher[2], J. Vandenbroucke[2], V. Vasileiou[27], G. Vianello[2,64], N. Vilchez[12,13], V. Vitale[50,66], A. P. Waite[2], E. Wallace[17], P. Wang[2], B. L. Winer[60], M. T. Wolff[33], D. L. Wood[67,68], K. S. Wood[33], Z. Yang[28,29], AND S. Zimmer[28,29]

[1] Deutsches Elektronen Synchrotron DESY, D-15738 Zeuthen, Germany
[2] W. W. Hansen Experimental Physics Laboratory, Kavli Institute for Particle Astrophysics and Cosmology, Department of Physics and SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94305, USA; monzani@slac.stanford.edu
[3] Istituto Nazionale di Fisica Nucleare, Sezione di Perugia, I-06123 Perugia, Italy
[4] Dipartimento di Fisica, Università degli Studi di Perugia, I-06123 Perugia, Italy
[5] Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, I-56127 Pisa, Italy
[6] Laboratoire AIM, CEA-IRFU/CNRS/Université Paris Diderot, Service d'Astrophysique, CEA Saclay, 91191 Gif sur Yvette, France
[7] Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, I-34127 Trieste, Italy
[8] Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy
[9] Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy
[10] Dipartimento di Fisica "G. Galilei," Università di Padova, I-35131 Padova, Italy
[11] Santa Cruz Institute for Particle Physics, Department of Physics and Department of Astronomy and Astrophysics, University of California at Santa Cruz, Santa Cruz, CA 95064, USA
[12] CNRS, IRAP, F-31028 Toulouse Cedex 4, France; vilchez@cesr.fr
[13] GAHEC, Université de Toulouse, UPS-OMP, IRAP, Toulouse, France
[14] Dipartimento di Fisica "M. Merlin" dell'Università e del Politecnico di Bari, I-70126 Bari, Italy
[15] Istituto Nazionale di Fisica Nucleare, Sezione di Bari, 70126 Bari, Italy
[16] Laboratoire Leprince-Ringuet, École polytechnique, CNRS/IN2P3, Palaiseau, France
[17] Department of Physics, University of Washington, Seattle, WA 98195-1560, USA
[18] Institut de Ciències de l'Espai (IEEE-CSIC), Campus UAB, 08193 Barcelona, Spain
[19] INAF-Istituto di Astrofisica Spaziale e Fisica Cosmica, I-20133 Milano, Italy; salvetti@lambrate.inaf.it
[20] Agenzia Spaziale Italiana (ASI) Science Data Center, I-00044 Frascati (Roma), Italy
[21] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; elizabeth.c.ferrara@nasa.gov
[22] Center for Research and Exploration in Space Science and Technology (CRESST) and NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[23] Department of Physics and Center for Space Sciences and Technology, University of Maryland Baltimore County, Baltimore, MD 21250, USA
[24] Artep Inc., 2922 Excelsior Springs Court, Ellicott City, MD 21042, USA
[25] National Academy of Sciences, Washington, DC 20001, USA
[26] ASI Science Data Center, I-00044 Frascati (Roma), Italy
[27] Laboratoire Univers et Particules de Montpellier, Université Montpellier 2, CNRS/IN2P3, Montpellier, France
[28] Department of Physics, Stockholm University, AlbaNova, SE-106 91 Stockholm, Sweden
[29] The Oskar Klein Centre for Cosmoparticle Physics, AlbaNova, SE-106 91 Stockholm, Sweden
[30] Dipartimento di Fisica, Università di Udine and Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Gruppo Collegato di Udine, I-33100 Udine, Italy
[31] Department of Physics and Department of Astronomy, University of Maryland, College Park, MD 20742, USA
[32] Istituto Universitario di Studi Superiori (IUSS), I-27100 Pavia, Italy
[33] Space Science Division, Naval Research Laboratory, Washington, DC 20375-5352, USA
[34] Department of Physical Sciences, Hiroshima University, Higashi-Hiroshima, Hiroshima 739-8526, Japan
[35] INAF Istituto di Radioastronomia, 40129 Bologna, Italy
[36] Max-Planck-Institut für Kernphysik, D-69029 Heidelberg, Germany
[37] Landessternwarte, Universität Heidelberg, Königstuhl, D 69117 Heidelberg, Germany

[38] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[39] Center for Space Plasma and Aeronomic Research (CSPAR), University of Alabama in Huntsville, Huntsville, AL 35899, USA
[40] Department of Astronomy, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
[41] School of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[42] Centre d'Études Nucléaires de Bordeaux Gradignan, IN2P3/CNRS, Université Bordeaux 1, BP120, F-33175 Gradignan Cedex, France
[43] Science Institute, University of Iceland, IS-107 Reykjavik, Iceland
[44] College of Science, Ibaraki University, 2-1-1, Bunkyo, Mito 310-8512, Japan
[45] Research Institute for Science and Engineering, Waseda University, 3-4-1, Okubo, Shinjuku, Tokyo 169-8555, Japan
[46] Istituto Nazionale di Fisica Nucleare, Sezioine di Torino, I-10125 Torino, Italy
[47] Yukawa Institute for Theoretical Physics, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan
[48] CNRS/IN2p3, Centre d'Études Nucléaires de Bordeaux Gradignan, Université Bordeaux 1, 33175 Gradignan, France
[49] Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK
[50] Istituto Nazionale di Fisica Nucleare, Sezione di Roma "Tor Vergata," I-00133 Roma, Italy
[51] Department of Physics, Boise State University, Boise, ID 83725, USA
[52] Hiroshima Astrophysical Science Center, Hiroshima University, Higashi-Hiroshima, Hiroshima 739-8526, Japan
[53] Institute of Space and Astronautical Science, JAXA, 3-1-1 Yoshinodai, Chuo-ku, Sagamihara, Kanagawa 252-5210, Japan
[54] Max-Planck Institut für extraterrestrische Physik, 85748 Garching, Germany
[55] Department of Physics and Astronomy, University of Denver, Denver, CO 80208, USA
[56] Max-Planck-Institut für Physik, D-80805 München, Germany
[57] Center for Earth Observing and Space Research, College of Science, George Mason University, Fairfax, VA 22030, USA
[58] Institut für Astro-und Teilchenphysik and Institut für Theoretische Physik, Leopold-Franzens-Universität Innsbruck, A-6020 Innsbruck, Austria
[59] NYCB Real-Time Computing Inc., Lattingtown, NY 11560-1025, USA
[60] Department of Physics, Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
[61] Department of Chemistry and Physics, Purdue University Calumet, Hammond, IN 46323-2094, USA
[62] Institut für Theoretische Physik and Astrophysik, Universität Würzburg, D-97074 Würzburg, Germany
[63] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[64] Consorzio Interuniversitario per la Fisica Spaziale (CIFS), I-10133 Torino, Italy
[65] INTEGRAL Science Data Centre, CH-1290 Versoix, Switzerland
[66] Dipartimento di Fisica, Università di Roma "Tor Vergata," I-00133 Roma, Italy
[67] Praxis Inc., Alexandria, VA 22303, USA

ABSTRACT

The *Fermi* Large Area Telescope (LAT) First Source Catalog (1FGL) provided spatial, spectral, and temporal properties for a large number of γ-ray sources using a uniform analysis method. After correlating with the most-complete catalogs of source types known to emit γ rays, 630 of these sources are "unassociated" (i.e., have no obvious counterparts at other wavelengths). Here, we employ two statistical analyses of the primary γ-ray characteristics for these unassociated sources in an effort to correlate their γ-ray properties with the active galactic nucleus (AGN) and pulsar populations in 1FGL. Based on the correlation results, we classify 221 AGN-like and 134 pulsar-like sources in the 1FGL unassociated sources. The results of these source "classifications" appear to match the expected source distributions, especially at high Galactic latitudes. While useful for planning future multiwavelength follow-up observations, these analyses use limited inputs, and their predictions should not be considered equivalent to "probable source classes" for these sources. We discuss multiwavelength results and catalog cross-correlations to date, and provide new source associations for 229 *Fermi*-LAT sources that had no association listed in the 1FGL catalog. By validating the source classifications against these new associations, we find that the new association matches the predicted source class in ~80% of the sources.

Key words: catalogs – galaxies: active – gamma rays: general – methods: statistical – pulsars: general

Online-only material: color figures, machine-readable tables

1. INTRODUCTION

Astrophysical sources of high-energy γ rays (photon energies above 10 MeV), although inherently interesting as tracers of energetic processes in the universe, have long been hard to identify. Only four of the 25 sources in the second COS-B catalog had identifications (Swanenburg et al. 1981), and over half the sources in the third EGRET catalog had no associations with known objects (Hartman et al. 1999). A principal reason

for the difficulty of finding counterparts to high-energy γ-ray sources has been the large positional errors in their measured locations, a result of the limited photon statistics and angular resolution of the γ-ray observations and the bright diffuse γ-ray emission from the Milky Way. In addition, a number of the COS-B and EGRET sources were determined to be spurious by follow-up analysis and observations.

A major step forward for detection and identification of high-energy γ-ray sources came when the *Gamma-ray Large Area Space Telescope* was launched on 2008 June 11. It began its scientific operations two months later, and shortly thereafter, it was renamed the *Fermi Gamma-ray Space Telescope*. Its primary instrument is the Large Area Telescope (LAT; Atwood et al. 2009), the successor to the Energetic Gamma-Ray Experiment Telescope (EGRET) on the *Compton Gamma-Ray Observatory* (Thompson et al. 1993). The LAT offers a major increase in

sensitivity over EGRET, allowing it to study the 100 MeV to ~300 GeV γ-ray sky in unprecedented detail.

The high sensitivity, improved angular resolution, and nearly uniform sky coverage of the LAT make it a powerful tool for detecting and characterizing large numbers of γ-ray sources. The *Fermi*-LAT First Source Catalog (1FGL; Abdo et al. 2010a) lists 1451 sources detected during the first 11 months of operation by the LAT, of which 821 were shown to be associated with at least one plausible counterpart. Of these, 698 were extragalactic (mostly active galactic nuclei or AGNs) and 123 were Galactic (mostly pulsars and supernova remnants (SNRs), but also pulsar wind nebulae (PWNe) and high-mass X-ray binaries). After the publication of the 1FGL catalog, the association panorama evolved very quickly with the release of the catalog of AGNs (1LAC; Abdo et al. 2010b) as well as a catalog of PWNe and SNRs (Ackermann et al. 2011).

Here, as a starting point for our multivariate classification strategy, we consider the entire original list of 630 1FGL sources that remain unassociated with plausible counterparts at other wavelengths. A plausible counterpart is a member of a known or likely γ-ray emitting class located close to the 95% uncertainty radius of a given 1FGL source, with an association confidence of 80% or higher (Abdo et al. 2010a). The 95% uncertainty radii for 1FGL source locations are typically 10′. While greatly improved over the degree-scale uncertainties of previous instruments, these position measurements are still inadequate to make firm identifications based solely on location.

We have taken a multi-pronged approach toward understanding these unassociated 1FGL sources, using all the available information about the γ-ray sources. Information about locations, spectra, and variability has been combined with properties of the established γ-ray source classes and multiwavelength counterpart searches.

Here we look in depth at the properties of the 1FGL unassociated sources, and investigate the implications of those characteristics. Specifically, this paper addresses five primary questions.

1. What do the γ-ray properties of the unassociated 1FGL sources reveal about these sources (Section 2)?
2. What does our understanding of the γ-ray properties of the associated sources suggest about the possible source class for each of the 1FGL unassociated sources (Section 3)?
3. What new associations or multiwavelength counterparts have been found beyond those from the first LAT catalog (Section 4)?
4. Do the new classifications properly predict sources that have been associated since the release of the 1FGL catalog (Section 5)?
5. What do the new classifications and associations imply about the existence of unknown new γ-ray source classes (Section 6)?

Although the 2FGL catalog (Nolan et al. 2012) was being developed in parallel with the present work, we focus on the 1FGL results, where some follow-up results are available for comparison with the methods of this work. Such follow-up observations for 2FGL have yet to be done.

2. GAMMA-RAY PROPERTIES OF UNASSOCIATED *FERMI*-LAT SOURCES

In the 1FGL catalog (hereafter "1FGL"; Abdo et al. 2010a), source identifications and associations were assigned through an objective procedure. For a source to be considered identified

in the 1FGL catalog, detection of periodic emission (pulsars or X-ray binaries) or variability correlated with observations at other wavelengths (blazars) was required. Additionally, measurement of an angular extent consistent with observations at other wavelengths was used to declare identifications for a few sources associated with SNRs and radio galaxies (Abdo et al. 2009a, 2010c, 2010d, 2010e). Associations were reported only for sources with positional correlations between LAT sources and members of plausible source classes (based on Bayesian probabilities of finding a source of a given type in a LAT error box). This automated procedure was based on a list of 32 catalogs that contain potential counterparts of LAT sources based either on prior knowledge about classes of high-energy γ-ray emitters or on theoretical expectations. In addition, it indicated coincident detections at radio frequencies and TeV energies, as well as positional coincidences with EGRET and *AGILE* sources.

In total 821 of the 1451 sources in the 1FGL catalog (56%) were associated with at least one counterpart by the automated procedure, with 779 being associated using the Bayesian method while 42 are spatially correlated with extended sources based on overlap of the error regions and source extents. From the simulations in 1FGL we expect that ~57 among the 821 sources (7%) are associated spuriously in 1FGL. We found the initial list of unassociated sources by simply extracting the list of 1FGL sources without any association from the 1FGL catalog. These sources are spread across the sky, with about 40% located within 10° of the Galactic plane.

Sources without firm identifications that are in regions of enhanced diffuse γ-ray emission along the Galactic plane or are near local interstellar cloud complexes (like Orion), sources that lie along the Galactic ridge ($300° < l < 60°$, $|b| < 1°$), and sources that are in regions with source densities great enough that their position error estimates overlap in the γ-ray data are called c-sources, as their 1FGL designator has a "c" appended to indicate "caution" or "confused region." The remainder of the unassociated sources did not have a "caution" designator in 1FGL, and here are called "non-c" sources.

The positions, variability, and spectral information given in the catalog provide an important starting point for the characterization of LAT unassociated sources. We can easily compare intrinsic properties of the 1FGL sources such as spectral index, curvature index, and flux in different energy bands for both associated and unassociated populations, potentially providing insight into the likely classes of the unassociated sources.

For the 1FGL catalog, the limiting flux for detecting a source with photon spectral index $\Gamma = 2.2$ and Test Statistic of 25 (TS = 2Δlog(likelihood); Mattox et al. 1996) varied across the sky by about a factor of five (see Figure 19 of 1FGL). This non-uniform flux limit is due to the non-uniform Galactic diffuse background and non-uniform exposure (mostly arising from the passage of the *Fermi* observatory through the South Atlantic Anomaly).

As discussed in 1FGL, when the variability and spectral curvature properties of *Fermi*-LAT sources are compared against each other, a clear separation is visible between bright sources with AGN associations and those with pulsar associations. In Figure 1 (top panel), pulsars lie in the lower right-hand quadrant and AGNs lie in the upper half. However, the two classes mix in the lower left-hand quadrant, making it difficult to distinguish between them. This region of parameter space is home to much of the unassociated source population (bottom panel). A closer look at these and other properties of the known sources gives
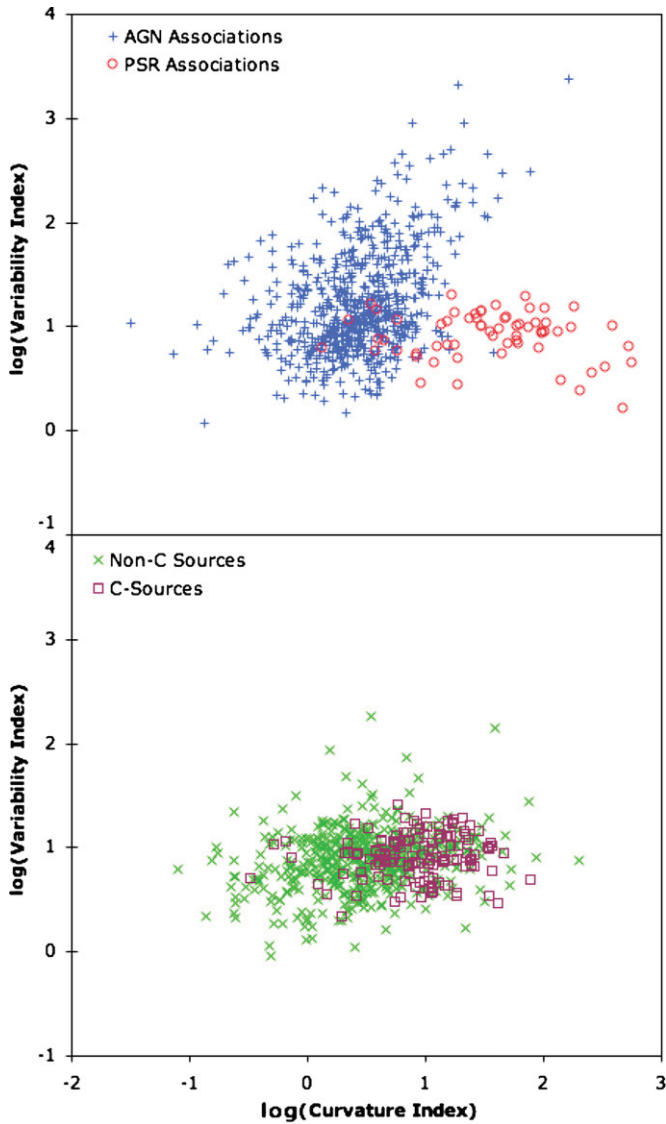
Figure 1. Comparison of the 1FGL Variability Index vs. Curvature Index for the associated sources (top panel) and unassociated sources (bottom). A separation between the AGN (crosses) and pulsar (circles) populations is evident. However the unassociated sources mainly lie in the region where those two populations overlap.

(A color version of this figure is available in the online journal.)

clues to methods of separating the two major types, allowing us to classify some of the unassociated sources as likely members of one of these two source types (Section 3).

2.1. Source Locations and Flux Distributions

The spatial distributions of the major source types (AGNs, pulsars, unassociated sources) are given in Table 1. It is clear that there is a significant excess of unassociated sources at low Galactic latitudes ($|b| < 10°$) where 63% of the detected sources have no formal counterparts, compared with only 36% unassociated at $|b| > 10°$.

Figure 2 shows the spatial distribution of LAT unassociated sources, with the positions of non-c sources shown as crosses and the c-source positions given by circles. As for the EGRET (3EG) catalog sources, the distribution is clearly not isotropic (Hartman et al. 1999). One consideration when interpreting the distribution of unassociated 1FGL sources is that a number of

Table 1
Spatial Distribution of Various Source Associations from the
1FGL and 1LAC Catalogs

Source Class	Sources at $|b| > 10°$	Sources at $|b| < 10°$	Ridge[a] Sources
Associated	**670**	**151**	**31**
AGN	642	51	1
Pulsars	16	47	11
SNRs/PWNe	1	45	19
Other	11	8	0
Unassociated	**373**	**257**	**88**
Non-c sources	354	139	0
c-sources	19	118	88

Notes. [a] Here, the Galactic ridge is defined as sources with $|b| < 1°$ and $|l| < 60°$. This value is a subset of the previous column of $|b| < 10°$ sources. Bold indicates the totals in each column for associated and unassociated sources. The unbolded values are subsets.

the remaining unassociated sources are in low Galactic latitude regions where catalogs of AGNs have limited or no coverage, reducing the fraction of AGN associations. If we bin the different source types by Galactic latitude (Figure 3), we see a clear absence of AGN associations in the central 10° of the Galaxy ($|b| < 5°$), while in the same region there is a spike in the number of unassociated sources.

The unassociated sources have an average flux of 3.1×10^{-9} photons cm^{-2} s^{-1} ($E > 1$ GeV), while the associated population averages are 5.5×10^{-8} photons cm^{-2} s^{-1} for pulsars and 2.7×10^{-9} photons cm^{-2} s^{-1} for AGNs.

In the Galactic plane a γ-ray source must be brighter than at high latitudes in order to be detected above the strong Galactic diffuse emission. Figure 4 (left panel) shows the 1FGL source flux distribution versus Galactic latitude for three longitude bands. It is clear that the Galactic plane ($|b| < 2°.5$) is dominated by Galactic diffuse emission, raising the flux detection threshold to $>5 \times 10^{-9}$ photons cm^{-2} s^{-1}. This is reflected in the average flux of the unassociated c-sources which at 8.2×10^{-9} photons cm^{-2} s^{-1} is significantly higher than that for the non-c unassociated source population (1.7×10^{-9} photons cm^{-2} s^{-1}). Outside the central region of the Galaxy, the flux threshold is lower than that shown in Figure 4.

As was the case for COS-B and EGRET, it is likely that a subset of the unassociated sources are spurious, resulting from an imperfect Galactic diffuse model. Such sources probably have very low significance, poor localization, and a spectral shape that mimics that of the Galactic diffuse emission itself. The c-sources in the 1FGL catalog are candidates to be sources of this type.

As discussed in Section 4.7 of 1FGL, the latitude distribution of the Galactic ridge ($300° < l < 60°$) unassociated sources shows a sharp narrow peak in the central degree ($|b| < 0°.5$) of the Galaxy (Figure 4, right). If this feature is not an artifact, and we assume these sources originate in a Galactic population, then the scale height for this population must be ~50 pc, to keep the average distance to the sources within the Galaxy. Such a scale height does not correspond to any known population of γ-ray sources, making it likely that a number of the sources in the Galactic ridge are spurious.

2.2. Spectral Properties

The 1FGL catalog provides spectral information that may be useful for distinguishing between different source classes.
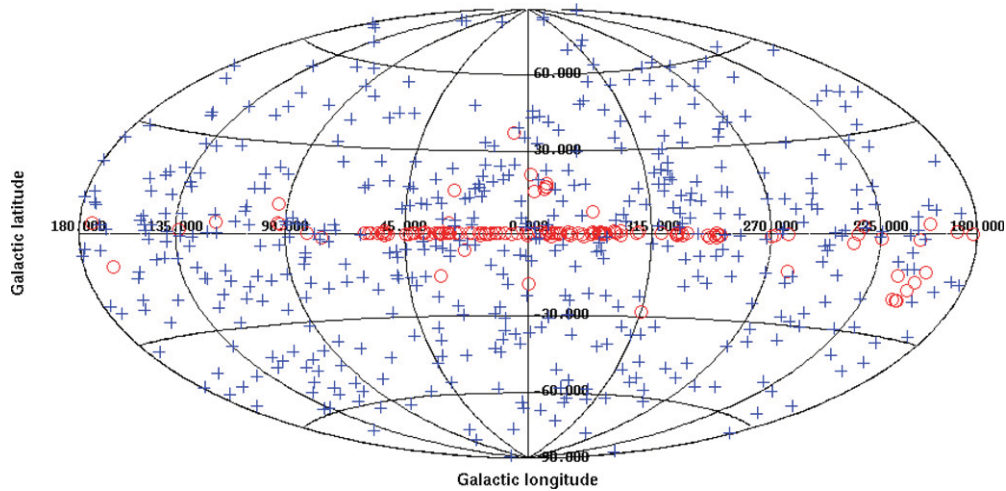
Figure 2. 1FGL sky map with the positions of the unassociated sources marked. Here, the non-c unassociated sources are indicated by crosses, the c-sources by circles.
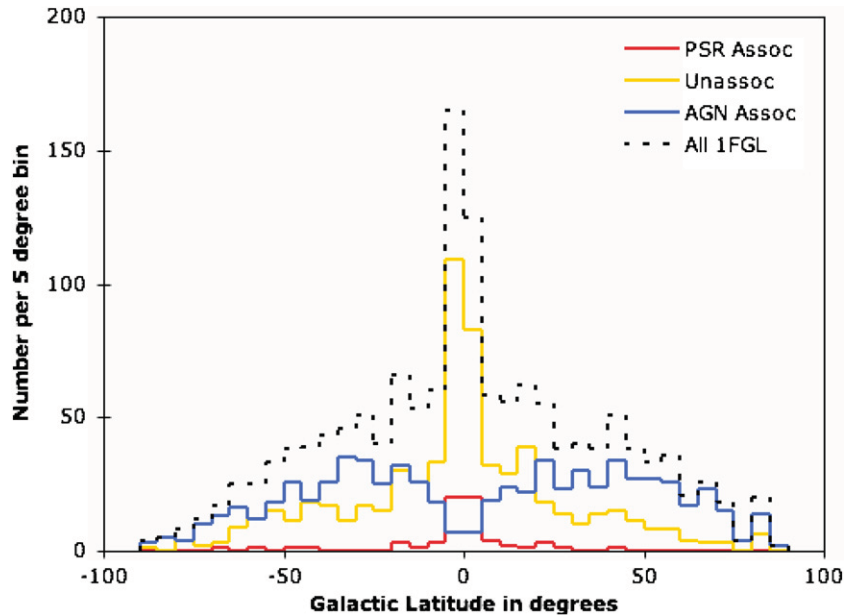(A color version of this figure is available in the online journal.)

Figure 3. Distribution of 1FGL source types by Galactic latitude. The sources associated with AGNs (blue line) show a clear deficit at low latitudes, while the same region hosts a large number of unassociated sources (yellow line) and identified pulsars (red line).
(A color version of this figure is available in the online journal.)

As part of the 1FGL analysis all sources were fit with a power-law spectral form and the spectral indices were included in the catalog. In addition, the catalog includes a "curvature index," which measures the deviation of the spectrum from the simple power-law form for each source. This means the curvature index is more a measure of the quality of the power-law spectral fit than of the intrinsic spectral shape. Figure 5 shows the distributions of the spectral index (top panel) and curvature index (middle panel) with respect to flux. Neither of these parameters appears to discriminate well between the AGN and pulsar populations. In addition, the relationship is nearly linear for the curvature index, indicating that this parameter is strongly correlated with flux. That is, fainter sources have relatively poorly measured spectra that cannot be measured to be significantly different from power laws. This means that faint γ-ray sources provide less discriminating information than bright sources.

The majority of γ-ray AGNs are blazars, which are relativistic jet sources with the jets directed toward Earth. An important property of blazars is their typical γ-ray spectral index, which offers some discrimination power between Flat Spectrum Radio Quasars (FSRQs) and BL Lac objects (Abdo et al. 2010f). The spectra of blazars in both of these sub-classes are typically well described as broken power laws in the LAT energy range, and the distributions of spectral indices for FSRQs and BL Lac objects are compatible with Gaussians (Abdo et al. 2009b, 2010b). However, because pulsar spectra are not well described by power laws, the spectral index of a power-law fit is not a good discriminator between pulsar and AGN classes.

As mentioned before, Figure 5 (middle) shows that the curvature index for pulsars is strongly correlated with flux. This is primarily because many of the pulsars detected in the 1FGL catalog are strong γ-ray sources, with brighter pulsars having a more significant spectral curvature than fainter pulsars.
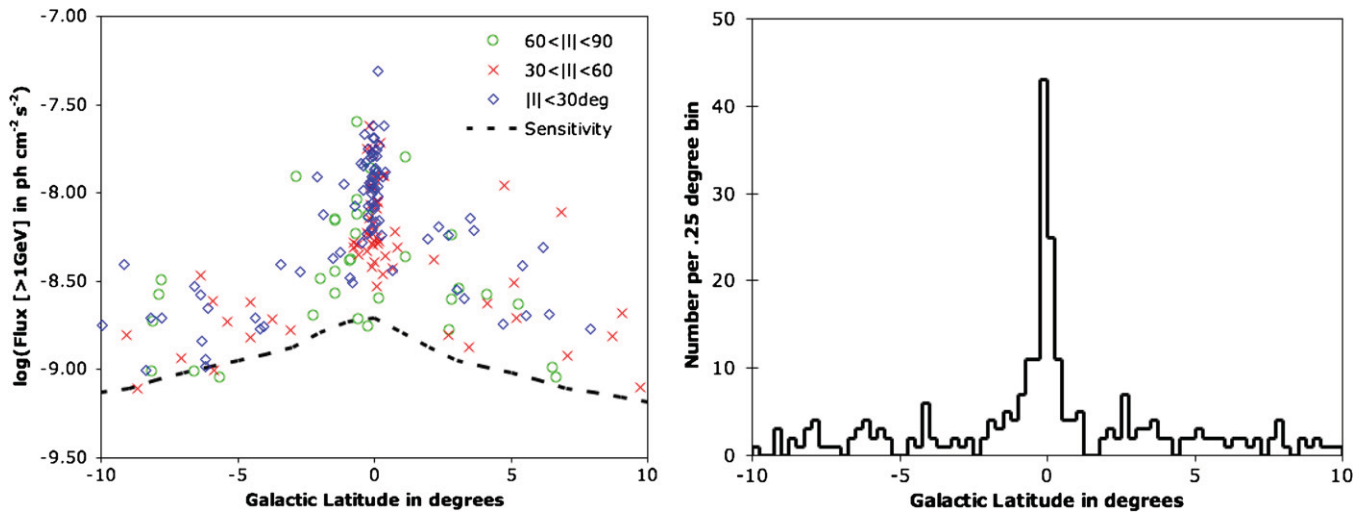
Figure 4. Distribution of unassociated sources in the Galactic ridge. Left: source flux ($E > 1$ GeV) for all 1FGL sources as a function of Galactic latitude in three longitude bands. The dashed line shows the threshold flux for detectability of a source with a power-law spectrum of photon spectral index $\Gamma = 2.2$ (from the 1FGL sensitivity map, at $|l| = 0$). An increase in minimum flux is clearly visible for sources near $|b| = 0°$. Right: unassociated source counts in 0.°25 bins. A sharp peak in the number of unassociated sources is visibly clustered along the central 0.°5 of Galactic latitude.

(A color version of this figure is available in the online journal.)
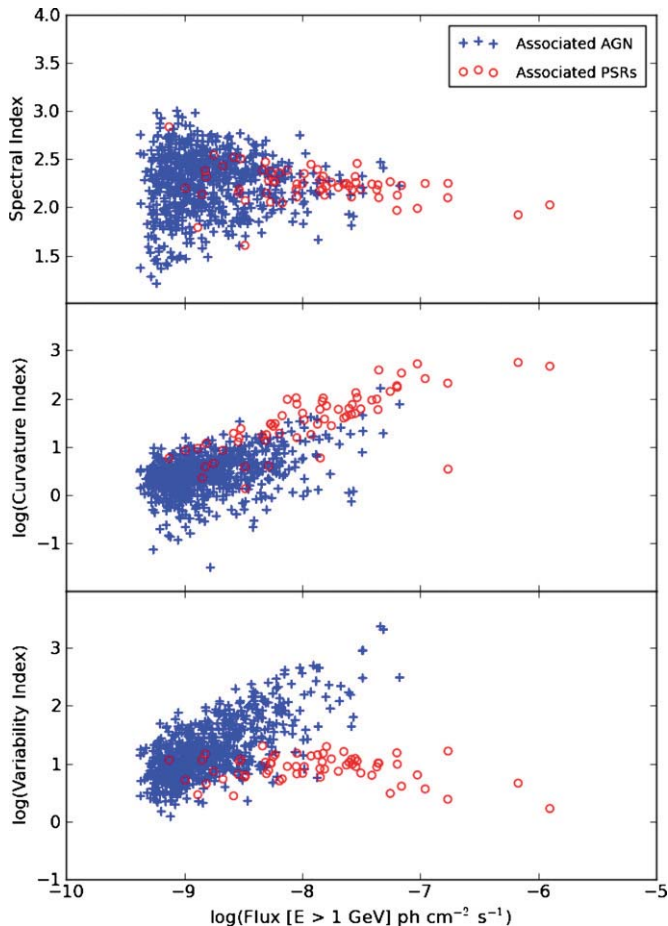

Figure 5. Distributions with respect to flux of the spectral index (top), curvature index (middle), and variability index (bottom) for the 1FGL associated and identified sources. It is clear that the curvature index is dependent on source flux for both AGN (crosses) and pulsar (circles) populations. The high flux pulsar with a low value for curvature index is the Crab pulsar.

(A color version of this figure is available in the online journal.)

Unfortunately, the broken power-law spectral forms of bright blazars (e.g., Ackermann et al. 2010) also have the effect of inducing a correlation between curvature index and flux for LAT blazars.

There are few sources with significant detections in all of the five spectral bands used to calculate the curvature index. Only 36 of the 630 unassociated sources are strongly enough detected to have flux measurements reported in each band in the 1FGL catalog, as this requires a TS >10 in each energy range. In contrast, 181 of the unassociated sources are detected in only a single band, with an additional 88 sources having upper limits in all the spectral bands (i.e., are only detected when data at all energies are combined).

2.3. Variability Properties

In the 1FGL catalog the variability index for each source was defined as the χ^2 of the deviations of eleven monthly (30 day) source flux measurements from the average source flux (Abdo et al. 2010a). While this value increases with flux for AGNs, it does not do so for the pulsars (Figure 5, bottom panel), making variability a much better discriminator between the two major classes. One property of blazars is that they are frequently significantly variable in γ rays (Abdo et al. 2010g). Their fluxes can vary up to a factor of five on timescales of a few hours and by a factor of 50 or more over several months. As a consequence, their characteristic variability can serve as a primary discriminator. This property has been used to turn some *Fermi*-LAT AGN associations into identifications due to their timing properties (Abdo et al. 2010g; Vandenbroucke et al. 2010). For variability to be a useful indicator, the timescale must be adapted to the source significance. Indeed, for a faint source, the variability needs to be tested on longer timescales than for a bright source. All sources in the 1FGL catalog were processed in the same way, regardless of flux. Thus, for the many 1FGL sources not bright enough to be significantly detected on monthly timescales, the 1FGL variability index is not a sensitive discriminator of variability.

Pulsars, on the other hand, are generally steady sources. Where variability has been seen in γ rays, it has been attributed
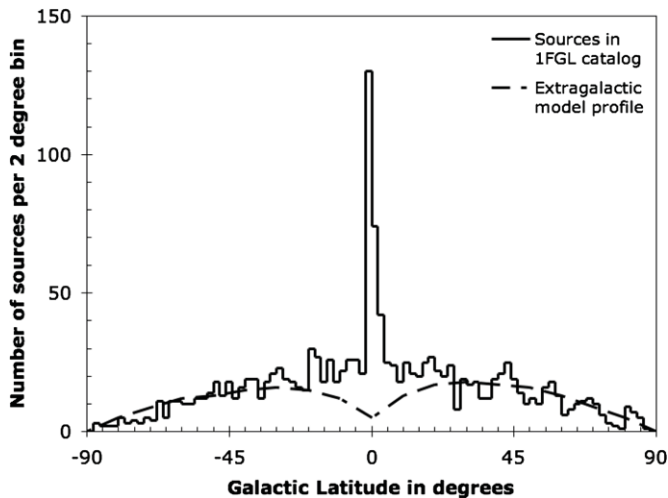
Figure 6. Latitude profile of the 1FGL sources and the extragalactic source model profile (dashed line).

Table 2
Expected versus Observed Source Distribution

Source Types	Sources at $|b| > 10°$	Sources at $|b| < 10°$	Totals
Total detected	1043 (71.9%)	408 (28.1%)	1451
Associated	670	151	821
Unassociated	373	257	630
Extragalactic			
Total from log N–log S	972	88	1060 (73.1%)
Associated	650	51	701
Not associated in 1FGL	322	37	359
New Associations	150	27	177
New Classifications	154	67	221
log N–log S comparison	−18	+57	+39
Galactic			
Total from log N-log S	71	320	391 (26.9%)
Associated	20	100	120
Not associated in 1FGL	51	220	271
New Associations	27	25	52
New Classifications	31	103	134
log N–log S comparison	+7	−92	−85

Note. Results from log N–log S analysis, applied to the 1FGL source list.

to flares in the nebular contribution of a PWN, rather than to the pulsar itself (Tavani et al. 2011; Abdo et al. 2011). This flux stability places pulsars in extreme opposition to AGNs in the γ-ray regime. Essentially any significant detection of variability in an unassociated source is enough to make a pulsar classification extremely unlikely.

In the 1FGL catalog, 241 sources were found to be variable at a formal confidence level of 99% (variability index >23.21). Of these, two are HMXBs, 221 are AGNs, and 18 are unassociated. Variability in bright sources is easier to detect as the source flux is typically above the sensitivity threshold in each monthly bin. For the lower-flux unassociated sources, however, we need a method to improve the detection of variability. Using the fractional variability (discussed in Section 3.1) is one such method.

2.4. Comparison with Source Modeling

We can also examine what source distributions we might expect given the populations of source types in 1FGL. We do this by first estimating the total number of detected AGNs in 1FGL, which we derive from a model population. To quantify the total number of AGNs, we model the population and then apply that model to the 1FGL catalog.

We use the *Fermi*-LAT log N–log S distribution (the distribution of the number N of sources detectable at a given sensitivity S) for AGNs (Abdo et al. 2010f), and the 1FGL sensitivity map (Figure 19 of Abdo et al. 2010a) to generate a Galactic map that contains the number of the expected AGN at each position in the sky. Summing these results over Galactic longitude we obtain the AGN latitude profile shown in Figure 6. Integrating the AGN model allows us to estimate the number of expected AGNs in the sky. By subtracting the number of AGN found by the model from the sources in the 1FGL catalog we obtain the number of Galactic sources in and out of the plane. This gives the Galactic source estimates for the unassociated source list.

Table 2 compares the 1FGL source counts with this model for low- and high-latitude regions. It is clear that the group of sources that is most difficult to associate are those of Galactic origin at low Galactic latitudes. This is likely due to the presence of a population of spurious sources in that region in the 1FGL catalog.

At high Galactic latitudes, pulsars are the second most numerous class of identified γ-ray sources; most of those are millisecond pulsars (MSPs). From the set of 1FGL pulsars and the new pulsar associations discussed herein, we find that more than a third of the γ-ray pulsars known to date are MSPs. If we then assume that ~50% of the 271 unassociated sources that are expected to be Galactic sources are pulsars (based on the fraction of Galactic sources that are pulsars as given in Table 1), and one-third of those pulsars are MSPs, we find that we expect 45 new MSPs in the 1FGL unassociated source list. Of the 31 new MSPs discovered to date (Section 4.2), 28 are at high Galactic latitudes, suggesting that an expected number of 45 is not unreasonable.

At low Galactic latitudes the source content is more diverse, with half the sources being compact objects (pulsars and X-ray binaries) and nearly half being extended sources (SNRs and PWNe). If the unassociated sources have a similar distribution, then there will be ~100 pulsars and ~100 SNRs/PWNe.

3. CLASSIFICATION OF UNASSOCIATED SOURCES

The spatial, spectral, and variability properties discussed in Section 2 provide a framework that allows us to try to predict the expected source classes for the sources that remain unassociated. This is done by using the properties of the associated sources to define a model that describes the distributions and correlations between measured properties of the γ-ray behavior of each source class. This model is then compared to the γ-ray properties of each unassociated source. Generating the model requires an associated source parent population with enough members to describe the behavior well. For this reason, we have focused only on AGNs and pulsars as the input source populations.

To create a model, it is necessary to use γ-ray properties that are clearly different between the parent populations. For 1FGL, the best parameters are the spectral index, curvature index, and variability index, as well as hardness ratios between the different spectral bands. In addition, it is important that the properties used to generate the model not be related to source significance, as this will bias the results. However, as discussed in Section 2.1, the curvature index appears to be dependent on

the source flux, and thus is not a good indicator of the parent population. Also, the spectral index and hardness ratios are both spectral indicators, so they overlap in functionality. Since the hardness ratios provide more information about spectral shape than the spectral index, they are preferred for this analysis.

To generate valid classifications, we must first define new parameters that allow intrinsic properties to be compared rather than relative fluxes. With the new parameters in hand, we can generate classification predictions using multiple methods and compare these predictions to each other. However, the results from these techniques assume that the training samples and test samples have the same distributions of intrinsic properties. This is not true for the 1FGL unassociated sources, as they are more frequently found in the plane of the Galaxy with elevated background levels and in confused regions. To help compensate for this difference, we will validate the results against an independent set of classified sources.

3.1. Improving Source Type Discriminators

To mitigate the effect of low fluxes on the determination of the band spectra, it is necessary to define additional comparative parameters that remove the significance dependency. In this case, the 1FGL catalog provides a set of fluxes in five bands for each source from which we can find hardness ratios. To get a normalized quantity the hardness ratios are constructed as:

$$HR_{ij} = (EnergyFlux_j - EnergyFlux_i)/(EnergyFlux_j + EnergyFlux_i). \quad (1)$$

This quantity will always be between -1 and 1; -1 for a very soft source ([high]EnergyFlux$_j = 0$) and $+1$ for a very hard source ([low]EnergyFlux$_i = 0$). Here, energy flux in $\log(E)$ units (i.e., νF_ν) is used instead of photon flux because the definition works well only when the quantities are of the same order. This is true for the energy fluxes (because the spectra are not too far from an E^{-2} power law) but not for photon fluxes.

It is also possible to define a quantity that discriminates curvature by combining two hardness ratios, preferably from bands with a high number of detected sources. Here we use (HR$_{23}$–HR$_{34}$), where bands 2, 3, and 4 are for 0.3–1, 1–3, and 3–10 GeV respectively. This hardness ratio difference, or curvature, is positive for spectra curved downward in νF_ν (like pulsars), zero for power laws, and negative for spectra curved upward (with a strong high-energy component).

To remove the source significance dependency for variability, we use the fractional variability (as defined in Equation 5 of Abdo et al. 2010a) instead of the variability index. The fractional variability is

$$FracVar = \sqrt{\frac{\sum_i (Flux_i - Flux_{av})^2}{(N_{int} - 1)Flux_{av}^2} - \frac{\sum_i \sigma_i^2}{N_{int}Flux_{av}^2} - f_{rel}}, \quad (2)$$

where N_{int} is the number of time intervals (11 in 1FGL), σ_i is the statistical uncertainty in F_i, and f_{rel} is an estimate of the systematic uncertainty on the flux for each interval. (Here we use 3% as in 1FGL.) For some 1FGL sources the quantity inside the square root is negative. Those sources are assigned a fractional variability of 2%. Figure 7 shows fractional variability versus the curvature (HR$_{23}$–HR$_{34}$) for both the associated (top panel) and unassociated (bottom panel) 1FGL sources.

To allow the largest possible sample when performing the classification, we use the actual best-fit values in calculating both of these quantities, even when the 1FGL analysis reported
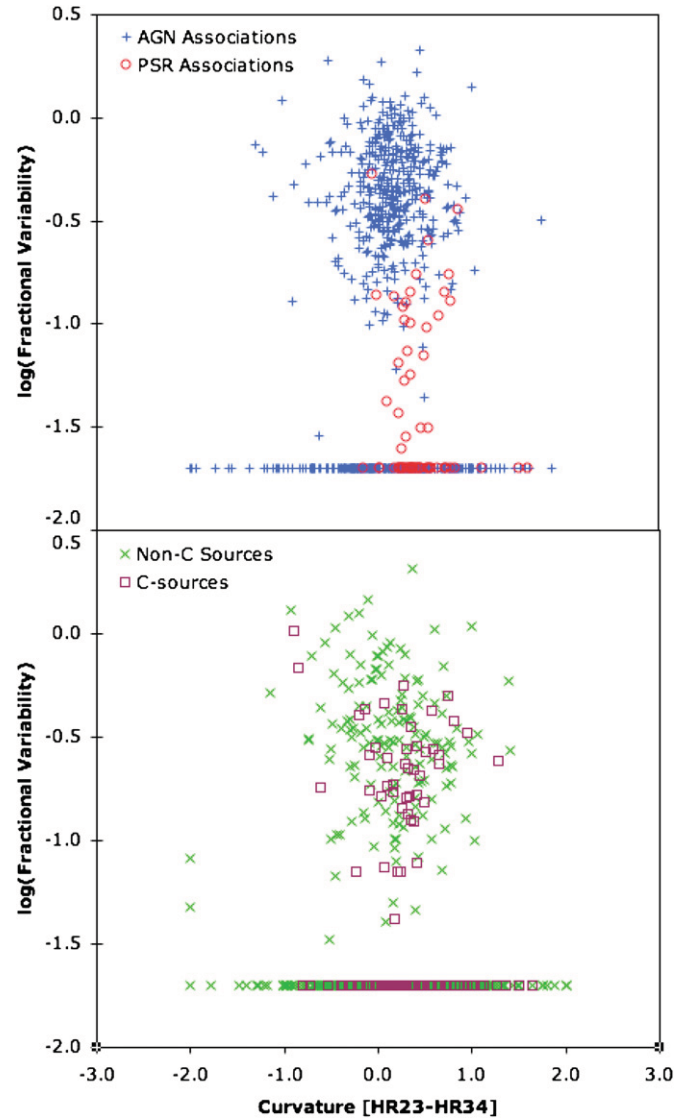

Figure 7. Fractional variability vs. hardness ratio difference. Top: the 1FGL associated AGNs (blue crosses) and pulsars (red circles). Bottom: the 1FGL non-c unassociated sources (green crosses) and the c-sources (purple squares).

(A color version of this figure is available in the online journal.)

only 2σ upper limits. For such sources, the variable is not well constrained. However, it will contribute to the distribution used to determine the classifications and may provide a small amount of additional discriminating power.

3.2. Classification Using Classification Trees

We have implemented two different data mining techniques to determine likely source classifications for the 1FGL unassociated sources: Classification Trees (CT; this section) and Logistic Regression (LR; Section 3.3). Both techniques use identified objects to build up a classification analysis which provides the probability for an unidentified source to belong to a given class. We applied these techniques to the sources in 1FGL to provide a set of classification probabilities for each unidentified source.

Classification Trees are a well-established class of algorithms in the general framework of data mining and machine learning (Breiman et al. 1984). The general principle of machine learning is to train an algorithm to predict membership of cases or objects to the classes of a dependent variable from their measurements

on one or more input variables. The advantage of this class of algorithms is the ability to produce a unique predictor parameter that takes into account several input quantities simultaneously. Other well-known flavors of machine learning algorithms include artificial neural networks, support vector machines, and Bayesian networks.

The purpose of analyses via tree-building algorithms is to determine a set of *if–then* logical conditions (called tree-nodes) that permit accurate prediction or classification of cases: The training procedure generates a tree in which each decision node contains a test on some input variable's value. The trees in this analysis are built through a process known as binary recursive partitioning, which is an iterative process of splitting the data into partitions. Initially all the records in the training set are assigned to a single class: The algorithm then tries breaking up the data, using every possible binary split on every field. The peculiar advantage of Classification Trees for our specific application is their flexibility in handling sparse or uneven distributions.

The specific flavor of algorithm used in this case was Adaptive Boosting, which reweights the importance of input sources at each step of the classification. The training of boosted decision trees is a recursive procedure, whereupon the weights of each incorrectly classified example are increased at each step, so that the new classifier focuses more on those examples. The output of such a procedure is actually a collection of trees, all grown from the same input sample: The selection will be done on a majority vote on the result of several decision trees (200 in the present case). In the following text, we will always refer to a single Classification Tree for simplicity, even though the real classifier is a much more complex object.

The training and application of Classification Trees to this specific analysis has been performed using TMine, an interactive software tool for processing complex classification analyses developed within the *Fermi*-LAT collaboration (Drlica-Wagner & Charles 2011). TMine is based on ROOT (Brun & Rademakers 1997), a very popular data analysis framework for high energy physics experiments. For the processing and parallel evaluation of multivariate classification algorithms, TMine utilizes the ROOT Toolkit for Multivariate Analysis (Hocker et al. 2007).

3.2.1. Selection of the Relevant Training Variables

The first step of the CT analysis is to select a sample of data to build the predictor variable. We decided to focus on the two most abundant classes of objects in the *Fermi* catalog, AGNs and pulsars, and to train a Classification Tree to discriminate between them. We trained a predictor using all the associated AGNs and pulsars in the 1FGL catalog (Abdo et al. 2010a). Because of the spectral similarities with pulsars, potential associations for sources near SNRs have been compounded with the pulsars sample. The output of this training process is a parameter (the predictor) that describes the probability for a new source to be an AGN.

Choosing the most appropriate set of variables for training the Classification Tree is a very delicate step in the analysis. It is extremely important to ensure that the selected variables are not dependent on the flux, the location, or the significance of the source: This check can be accomplished by comparing the distribution of the various parameters for associated and unassociated sources. Physical considerations about the γ-ray properties of each source class also guided us in the choice of the

Table 3
Ranking of the Training Variables for the Classification Tree

Variable	Importance
Fractional variability	21.9%
Hardness$_{45}$	16.0%
Hardness$_{23}$	15.8%
Spectral index	13.0%
Hardness$_{12}$	12.7%
Hardness$_{34}$	11.8%
Curvature	8.8%

Note. List of the training variables for the Classification Tree: each variable is ranked according to its relevance in the discrimination process, as computed by the CT algorithm.

most effective variables for discriminating AGNs from pulsars (Section 3.1).

After exploring most of the parameters in the 1FGL catalog, we selected a set of variables that includes: the curvature (HR_{23}–HR_{34}), the spectral index, and the fractional variability of each source, plus the hardness ratios for the five energy bands in the catalog. Table 3 ranks the relative importance of the different variables at distinguishing AGNs from pulsars: The weight of each variable was computed by the Classification Tree algorithm during the training process.

As described in Section 3.1, we used for training the actual best-fit values for each variable, even when the 1FGL analysis reported only upper limits. In case of faint sources that are not detected in one of the energy bands, some of the hardness ratios will be very close to -1 or 1: this is when the ability of Classification Trees in handling sparse distributions is particularly useful. Similar considerations apply to the fractional variability, which is very close to zero for pulsars, but varies for AGNs.

We chose not to use the Galactic latitude as input to the CT in order to avoid biasing our selection against AGNs situated along the Galactic plane and pulsars (especially MSPs) situated at high Galactic latitude. Furthermore, this choice gives us the opportunity to use the latitude distributions of the different populations as a cross check of our result. (The new pulsar candidates should be mainly distributed along the Galactic plane, while the AGN candidates should be isotropically distributed.)

While training the Classification Tree, 30% of known AGNs and pulsars, randomly selected from the input sample, were kept aside for cross-validation of the method. Such cross-validation was performed by comparing the predictor distributions for the training and testing samples via a Kolmogorov–Smirnov test. The result of the test provides a 93% probability for the AGN distributions and a 47% probability for the pulsar distributions, which make the test fully satisfactory.

It must also be noted that the sources associated with a different class than AGN or pulsar have been excluded from this training procedure, for a total of 24 sources. We cannot treat these 24 sources uniformly as "background," because of the smallness of their sample and the diversity of their spectral properties. However, it is possible to estimate the contamination to the candidate AGN and pulsar samples deriving from the likely presence of these "other" sources in the unassociated sample.

3.2.2. Output of the Classification Procedure

The second step of the analysis consists of deriving the predictor variable for unassociated sources by applying the Classification Tree that was trained in the previous step. The
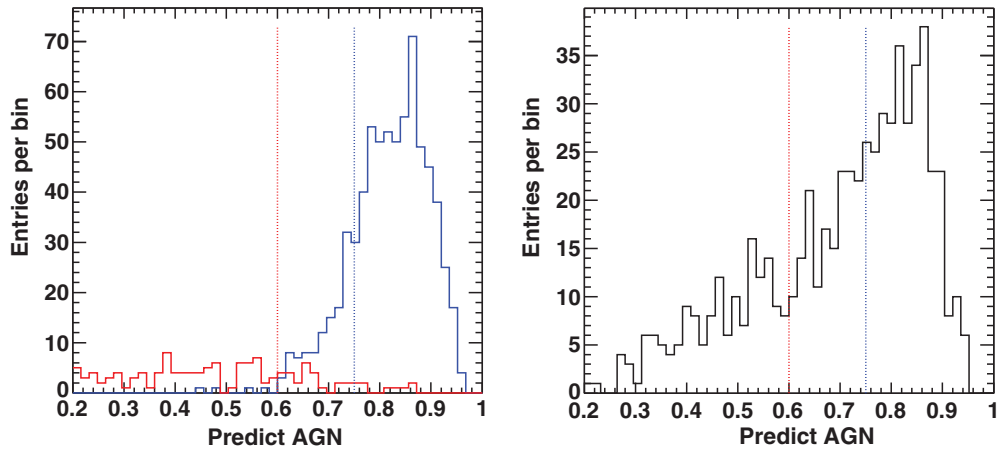
Figure 8. Distribution of the Classification Tree predictor. Vertical lines indicate the value of the thresholds we set to identify AGN candidates (Predictor >0.75) and pulsar candidates (Predictor <0.6). Left: sources of the 1FGL catalog identified as pulsar (red) and AGN (blue). Right: distribution of the predictor for unassociated sources.

(A color version of this figure is available in the online journal.)

Table 4
Summary of γ-ray Properties and Classification Results

1FGL Name	R.A.	Decl.	l	b	Var[a]	BSL[b]	Class Tree		Logistic Reg.		Combined	Radio	Radio	X-Ray	TeV
	(J2000)	(J2000)					Predict	Class	Predict	Class	Class	Imaging	Timing		
J0000.8+6600c	0.209	66.002	117.812	3.635			0.64		0.08						
J0001.9−4158	0.483	−41.982	334.023	−72.029			0.84	AGN	0.00	AGN	AGN	T			
J0003.1+6227	0.798	62.459	117.388	0.108			0.77	AGN	0.06	AGN	AGN	T			
J0004.3+2207	1.081	22.123	108.757	−39.448			0.87	AGN	0.00	AGN	AGN	T			
J0005.1+6829	1.283	68.488	118.689	5.999			0.78	AGN	0.00	AGN	AGN				
J0006.9+4652	1.746	46.882	115.082	−15.311			0.87	AGN	0.00	AGN	AGN	T			
J0008.3+1452	2.084	14.882	107.655	−46.708			0.77	AGN	0.00	AGN	AGN				
J0009.1+5031	2.289	50.520	116.089	−11.789			0.85	AGN	0.00	AGN	AGN				
J0016.6+1706	4.154	17.108	111.135	−44.964			0.87	AGN	0.00	AGN	AGN				
J0017.7−0019	4.429	−0.326	104.735	−62.001			0.84	AGN	0.00	AGN	AGN				

Notes. Summary of the γ-ray properties of the 1FGL unassociated sources. The table includes flags for variability, predictor values for both Classification Tree and Logistic Regression analyses, combined classification, and flags indicating what type of follow-up observations are recommended.
[a] T indicates the source was found to be variable in the 1FGL catalog analysis (Abdo et al. 2010a).
[b] T indicates the source was reported in the *Fermi*-LAT Bright Source List (Abdo et al. 2009e).

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

resulting predictor is a parameter that describes the probability that each of the unassociated sources is an AGN. The predictor is included in Table 4, which lists all 630 unassociated *Fermi*-LAT 1FGL sources and combines results for all the analyses discussed within this paper.

Figure 8 shows the distribution of the predictor for the 1FGL associated sources used in the training of the tree (left panel) and the distribution of the predictor for the unassociated sources (right panel). The global shapes of the two distributions are clearly different, with an apparent excess of pulsar-like sources among the unassociated sources when compared with the associated source distribution. This may be due to the presumably different fractions of AGNs and pulsars in the associated and unassociated samples, or there may be an additional contributing component. Nevertheless, the distribution of associated sources clearly shows that we can select a set of AGN and pulsar candidates with high confidence, when choosing the appropriate fiducial regions.

We set two fiducial thresholds: All the sources with a predictor greater than 0.75 are classified as AGN candidates while all the sources with a predictor smaller than 0.6 are classified as pulsar candidates. All the sources with an intermediate value of the

predictor remain unclassified after the CT analysis. The choice of these boundaries is optimized for an efficiency of 80% for the two source classes in order to keep the misclassification fraction under 2% (the misclassified fraction for a certain efficiency is determined by the width of the predictor distribution). Here, 80% of AGN associations in 1FGL have a predictor greater than 0.75 and 80% of pulsars have a predictor smaller than 0.6.

In this case, the extrapolation from the value of the predictor to the probability of class membership was performed empirically from the combined input sample (which includes both the training and testing samples). The expected misclassification fraction in each class was also evaluated with the same method. This analysis was repeated using the training and testing samples separately and yielded identical results. A more complex study used the area under the Receiver-Operating-Characteristic (ROC) curve that is obtained by plotting all combinations of true positives and the proportion of false negatives generated by varying the decision threshold. This study provided similar extrapolation results, but more optimistic misclassification fractions: We therefore decided to rely on the more conservative misclassification estimation provided by the combined input sample.

The predictor distribution for the 24 sources that were not used during the training procedure can be used to estimate the further contamination from these sources to the AGN and pulsar candidate distributions. According to the CT analysis, those 24 sources are equally distributed between AGN-like objects, pulsar-like objects, and still unclassified objects. Moreover, it is reasonable to assume that those sources will not be overrepresented in the unassociated population compared with the associated one. Therefore, we expect that up to 2% of the newly classified AGN candidates and up to 4% of the newly classified pulsar candidates will indeed belong to one of the "other" classes (galaxies, globular clusters, supernova remnants, etc.).

3.3. Classification Using Logistic Regression

Another approach to assign likely classifications for the 1FGL unassociated sources is the LR analysis method (Hosmer & Lemeshow 2000). Unlike the CT analysis, LR allows us to quantify the probability of correct classification based on fitting a model form to the data.

LR is part of a class of generalized linear models and is one of the simplest data mining techniques. LR forms a multivariate relation between a dependent variable that can only take values from 0 to 1 and several independent variables. When the dependent variable has only two possible assignment categories, the simplicity of the LR method can be a benefit over other discriminant analyses.

In our case, the dependent variable is a binary variable that represents the classification of given 1FGL unassociated source. Quantitatively, the relationship between the classification and its dependence on several variables can be expressed as

$$P = \frac{1}{(1 + e^{-z})}, \qquad (3)$$

where P is the probability of the classification and z can be defined as a linear combination:

$$z = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n, \qquad (4)$$

where b_0 is the intercept of the model, the b_i ($i = 0, 1, 2, \ldots, n$) are the slope coefficients of the LR model, and the x_i ($i = 0, 1, 2, \ldots, n$) are the independent variables. Therefore, LR evaluates the probability of association with a particular class of sources as a function of the independent variables (e.g., spectral shape or variability).

Much like linear regression, LR finds a "best-fitting" equation. However, the principles on which it is based are rather different. Instead of using a least-squared deviations criterion for the best fit, which is valid when the error terms are normally distributed, LR uses the general method of maximum-likelihood, which maximizes the probability of matching the associations in the training sample by optimizing the regression coefficients. As a result, the goodness of fit and overall significance statistics used in LR are different from those used in linear regression.

3.3.1. Selection of the Training Sample and the Predictor Variables

As LR is a supervised data mining technique, it must be trained on known objects in order to predict the membership of a new object to a given class on the basis of its observables. As with CT analysis, we trained the predictor using the pulsar and AGN associated sources in the 1FGL catalog (Abdo et al. 2010a). Like the CT analysis, the output of this training process

Table 5
List of the Predictor Variables for the LR Model

Variable	Coefficient	Standard Error	p-value
Intercept	-22.17	4.97	<0.001
Fractional variability	10.61	1.49	<0.001
Spectral index	11.30	2.47	<0.001
Hardness$_{23}$	-3.84	1.27	0.002
Hardness$_{34}$	8.14	1.53	<0.001
Hardness$_{45}$	3.72	0.76	<0.001
Hardness$_{12}$	$\ldots$	$\ldots$	0.242
glat	$\ldots$	$\ldots$	0.333
glon	$\ldots$	$\ldots$	0.144

Notes. Variables selected for the Logistic Regression analysis are listed at top. Those rejected are listed below the line.

is the probability that an unidentified source has characteristics more similar to an AGN than to a pulsar.

To evaluate the best predictor variables for the LR analysis, we used the likelihood ratio test, comparing the likelihoods of the models not including (null hypothesis) and including (alternative hypothesis) the predictor variable under examination. We started by using the fractional variability, the spectral index, the hardness ratios for the five energy bands in the catalog, and the position on the sky (i.e., the Galactic latitude and longitude). The value of the likelihood ratio test is the p-value and is useful in determining if a predictor variable is significant in distinguishing an AGN from a pulsar. If the p-value for a given predictor variable is smaller than the significance threshold α (0.05) then the predictor variable is included in the multivariate LR model. We did not include the curvature value (HR$_{23}$–HR$_{34}$) in this evaluation because the LR analysis does not work well with predictor variables that are linearly dependent on other predictor values.

We then calculated the significance of each predictor variable to find the resulting LR coefficients. The list of the LR predictor variables with the relative values of the maximum likelihood ratios can be found in Table 5.

While AGNs are isotropically distributed and pulsars are concentrated along the Galactic plane, we wanted to see whether our multivariate LR model was able to recognize this effect. The results indicate that Galactic latitude and longitude are not significant at the $\alpha = 0.05$ (5% significance) level. Moreover we find that also HR$_{12}$ is not highly significant in the LR analysis. In Table 5, those predictor variables selected for the LR model are above the line and those we did not select lie below the line.

3.3.2. Defining Thresholds

Next we derive the predictor variable for 1FGL unassociated sources by applying the trained classification analysis to those sources. Since the LR analysis used AGNs as primary source type, the output parameter (A) listed in Table 4 describes the probability that an unassociated source is an AGN. The probability that an unassociated source is a pulsar is $P = 1 - A$ (because we are modeling the behavior of AGNs as "opposite" of the behavior of the pulsars based on the predictor variables).

In principle, the dependent variable is a binary variable that represents the presence or absence of a particular class of objects. We could have selected "pulsars" and "non-pulsars" (e.g., all other 1FGL associated sources) to teach the model to recognize the new pulsars, and done similarly for the AGNs. We did not follow this approach because there are no source populations in 1FGL other than AGNs and pulsars with
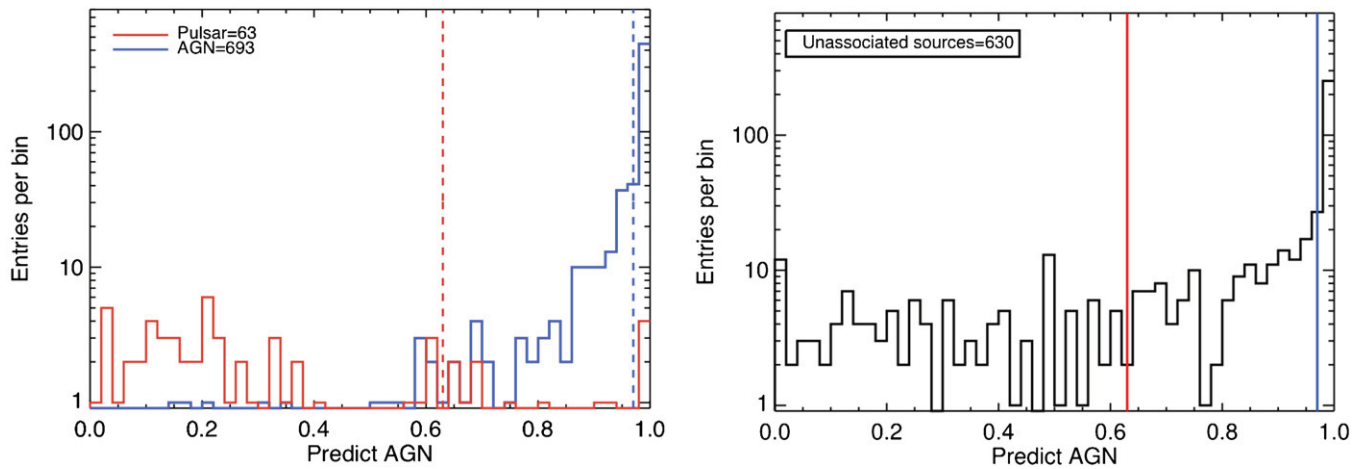
Figure 9. Distribution of the Logistic Regression predictor. Vertical lines indicate the value of the thresholds we set to identify pulsar candidates (Predictor <0.62) and AGN candidates (Predictor >0.98). Left: sources of the 1FGL catalog identified as pulsars (red) and AGNs (blue). Right: for 1FGL unassociated sources.
(A color version of this figure is available in the online journal.)

sufficient numbers to significantly affect the results. By focusing on "opposing" characteristics, we improve the efficiency of classifying new AGN or pulsar candidates.

As with the CT analysis we defined two threshold values, one to classify an AGN candidate (C_A) and one to classify a pulsar candidate (C_P). We chose these two thresholds by analyzing the ROC curves so that 80% of the AGN associations in 1FGL would have a predictor value greater than C_A and 80% of the pulsars would have a predictor value smaller than C_P, and to result in very low contamination. Using this principle we set C_A to 0.98 and C_P to 0.62. With these thresholds, only 1% of AGNs are misclassified as pulsars, while 3% of pulsars are classified as AGNs.

To estimate how accurately our predictive model performs in practice, we cross-validated using only the 756 pulsars and AGNs in the 1FGL catalog. We held out 75 sources to be the testing data set, and we used the remaining 681 for training. We repeated this procedure 10 times, using a different set of 75 test sources in each data set. At the end, this 10-fold cross-validation showed that the average testing efficiency rates for these threshold values are 75% for pulsars and 80% for AGNs, and that the average testing error rates (false positives) are very low, 0.5% for pulsars and 0.2% for AGNs. The 5% lower success rate for the pulsars is likely due to low statistics in the test samples.

If we apply the model to the 1FGL unassociated sources we find that 368 are classified as AGN candidates ($P > 0.98$), 122 are classified as pulsar candidates ($P < 0.62$), and 140 remain unclassified after the LR analysis. The distributions of 1FGL associated and unassociated sources as a function of the probability of being pulsars are shown in Figure 9. The thresholds for assigning pulsar candidates and AGN candidates are indicated in the figure. It is important to note that in order to meet the acceptance threshold of 80% of the known pulsars, we are including a large range of predictor values with very few pulsars. This may result in overpredicting the number of pulsars in unassociated sources.

3.4. Combining the Two Classification Methods

The two classification techniques gave somewhat different results. Of the 630 unassociated sources in 1FGL, both techniques agreed on the appropriate classification for 57.6% of the sources (363), while they gave conflicting classifications for 5.4% (34

Table 6
Comparison of Results for the Classification Techniques

	Classification		
	AGN	Pulsar	Unclassified
CT Totals	304	160	166
For $|b| > 10°$	244	33	96
For $|b| < 10°$	60	127	70
LR Totals	368	122	140
For $|b| > 10°$	276	39	58
For $|b| < 10°$	92	83	82
Combined	386	177	53
For $|b| > 10°$	300	50	22
For $|b| < 10°$	86	127	31
	LR Class		
CT Class	AGN	Pulsar	Unclassified
AGN	**269**	*2*	33
Pulsar	*32*	**72**	56
Unclassified	94	31	**41**

Notes. Summaries for both high- and low-Galactic latitude classification results for the Logistic Regression (LR) and Classification Tree (CT) techniques, as well as for the combined sample of classified sources (14 sources with conflicting classification are not included). In addition an inter-comparison of the two classification techniques is provided. Italics indicate conflicting results, while bold indicates agreement. All 630 1FGL sources are represented here.

sources). The remaining 253 sources were left unclassified by one or both techniques (see Table 6). Studies comparing these classification techniques (Perlich et al. 2003) have indicated that in data sets with good separability between the discriminating characteristics, the CT analysis should provide a more robust result. However, it is evident from the right panels of Figures 8 and 9 that the signal to noise for the unassociated sources does not provide such clear separability.

Since the purpose of this analysis is to provide candidate sources for follow-up multiwavelength studies, we use the positive results from both techniques to generate our candidate lists. From these two methods we can synthesize a final set of classifications for each source, where:

1. AGN candidates must be classified by at least one method, and the other method must not disagree (that is, not classify it as a pulsar).
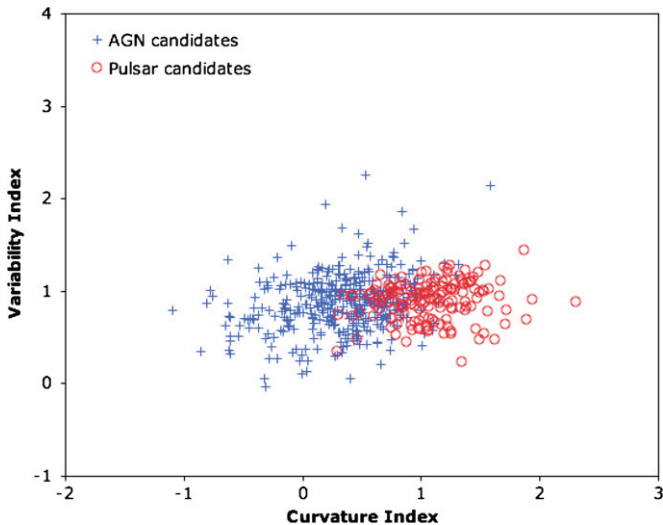
Figure 10. Variability index vs. curvature index for 1FGL unassociated sources classified as AGN (blue crosses) and pulsar candidates (red circles).
(A color version of this figure is available in the online journal.)

2. Pulsar candidates must be classified by at least one method, and the other method must not disagree (that is, not classify it as an AGN).
3. Unclassified sources are not classified by either method.
4. "Conflicting" sources are those that have been assigned opposite classifications (one AGN and one pulsar) by the two different methods.

Based on these definitions, there are 396 AGN candidates (269 are classified as AGNs by both methods), 159 pulsar candidates (72 classified as pulsars by both methods), 41 unclassified sources, and 34 conflicting sources in the 1FGL unassociated source list. Figure 10 shows the curvature-variability distribution of the newly classified AGN and pulsar candidates based on this synthesis of the two methods. We see that the unassociated sources have been separated into two populations with some overlap between them. Comparing this distribution to Figure 1 (top panel), we see that this separation follows the separation seen between the associated AGNs and pulsars.

4. RECENT ASSOCIATION EFFORTS

In order to validate the results of the classification methods, we must obtain an independent set of associations from those used to train the two methods. For this, we look to association efforts that have taken place since the release of the 1FGL catalog. This section will present the new associations, while the classification validation is reported in Section 5.

The associations listed in 1FGL were based on likely γ-ray-producing objects, i.e., those with energetic, non-thermal emission. The first LAT catalog of AGNs published shortly afterward found high-confidence AGN associations for 671 high Galactic latitude 1FGL sources, with an additional 155 LAT sources included in the low-latitude and lower confidence association lists (1LAC; Abdo et al. 2010b). The 1LAC association method was the same as for 1FGL, but the acceptance threshold for association was lower than for 1FGL.

For the unassociated sources in this paper, which by definition do not have plausible counterparts among the candidate catalogs used for comparison with the LAT sources, we must look beyond the obvious candidate source classes. Even with the improved source locations provided by the *Fermi*-LAT, however,

positional coincidence with a particular object is insufficient to claim an association.

If potential candidates can be found, then additional tests, based on spatial morphology, correlated variability, or physical modeling of multiwavelength properties offer the opportunity to expand the list of associations. X-ray, optical, or radio candidate counterparts all have better localizations than the γ-ray sources, so that a candidate in one of these wavelength bands can be matched with those in others. Also, most (if not all) of the catalogs and observations used here to find new associations are not complete surveys of the sky. Therefore the lack of an association for a 1FGL source does not mean that the source cannot be associated. In this section, we present only a preliminary report of the results from the many ongoing efforts to observe these fields in other wavebands.

4.1. Radio Searches for AGNs

The first step in searching for (or excluding) AGN counterparts of *Fermi*-LAT unassociated sources is to consult catalogs of radio sources. Almost all radio AGN candidates of possible interest are detected either in the NRAO VLA Sky Survey (NVSS; Condon et al. 1998) or the Sydney University Molonglo Sky Survey (SUMSS; Bock et al. 1999). NVSS covers the entire $\delta > -40°$ sky and provides interferometric flux density measurements at 1.4 GHz. SUMSS covers the remainder of the sky and offers interferometric measurements at 0.843 GHz.

In order to discover AGN counterparts, which are radio sources with compact, flat-spectrum cores, we follow the approach of the Cosmic Lens All Sky Survey (CLASS; Myers et al. 2003; Browne et al. 2003) and the Combined Radio All-Sky Targeted Eight GHz Survey (CRATES; Healey et al. 2007)—both of which have been shown (Abdo et al. 2009b, 2010b) to include substantial numbers of radio counterparts of LAT blazars—and pursue 8.4 GHz follow-up interferometry of blazar candidates. The *Fermi*-motivated VLA programs AH976 (Healey et al. 2009) and AH996 (in progress) obtained such data for several hundred sources, and ∼50 of these appear as "affiliations" (i.e., candidate counterparts for which quantitative association probabilities could not be computed) in the 1LAC catalog. One hundred and eight new associations, including the "figure of merit" value for each association (Abdo et al. 2010b), were determined by these VLA follow-up programs.

Recently, serendipitous radio identification surveys of 1FGL sources have been independently carried out using the recently released Australia Telescope 20 GHz radio source catalog (Murphy et al. 2010), which contains entries for 5890 sources observed at $\delta < 0°$. Cross-correlation between the 1FGL source list and the AT20G catalog has been performed by Mahony et al. (2010) and Ghirlanda et al. (2010). In particular, Mahony et al. (2010) find correlated radio sources for 233 1FGL sources and, based on Monte Carlo tests, they infer that 95% of these matches are genuine associations. While most of these radio detections are not identified or classified as specific object types, nine of the sources are considered likely to be AGNs (based on their properties at radio frequencies). Ghirlanda et al. (2010) obtain a similar number of matches with the AT20GHz catalog (∼230) and propose eight of the same sources as likely AGNs. All but one of these new associations were also found by the VLA observing program.

To date, radio observations of sources in the 1FGL unassociated sample have produced 109 new AGN associations. These new AGN associations are included in Table 7. In addition, the 1LAC catalog documented 57 other AGN associations with

Table 7
Summary of Follow-up Observations and Association Efforts

1FGL Name	R.A.	Decl.	l	b	EGRET[a]	Association			New[b]	Figure of[c]	Detections[d]	Reference
	(J2000)	(J2000)			Name	Identifier	Class		Observation	Merit		
J0000.8+6600c	0.209	66.002	117.812	3.635	. . .	. . .	. . .		. . .	. . .	. . .	. . .
J0001.9−4158	0.483	−41.982	334.023	−72.029	. . .	. . .	. . .		. . .	. . .	. . .	. . .
J0003.1+6227	0.798	62.459	117.388	0.108	. . .	. . .	. . .		. . .	. . .	*ROSAT*	15
J0004.3+2207	1.081	22.123	108.757	−39.448	. . .	. . .	. . .		. . .	. . .	. . .	. . .
J0005.1+6829	1.283	68.488	118.689	5.999	. . .	GB6 J0008+6837	AGN		VLA	0.804	*ROSAT*	15
J0006.9+4652	1.746	46.882	115.082	−15.311	. . .	. . .	. . .		. . .	. . .	*ROSAT*	15
J0008.3+1452	2.084	14.882	107.655	−46.708	. . .	RX J0008.0+1450	AGN		VLA	0.700	. . .	3
J0009.1+5031	2.289	50.520	116.089	−11.789	. . .	NVSS J000922+503028	AGN		VLA	0.941	. . .	. . .
J0016.6+1706	4.154	17.108	111.135	−44.964	. . .	GB6 J0015+1700	FSRQ		. . .	. . .	. . .	3
J0017.7−0019	4.429	−0.326	104.735	−62.001	. . .	S3 0013−00	FSRQ		VLA	0.570	. . .	3

Notes. Details of the follow-up observations used to determine the new associations. Where possible, a "figure of merit" was calculated for new AGN associations. Also listed are detections in other waveband that did not lead to new associations. References are given for associations based on other publications.
[a] Designator for an EGRET source that is positionally associated.
[b] Indicates what (if any) additional data were considered in determining the new association.
[c] A figure-of-merit (FOM) value for sources that were associated using the technique defined in the 1LAC paper (Abdo et al. 2010b); an ellipsis indicates an affiliation that does not have enough information to calculate an FOM.
[d] Waveband or catalog with association candidates.
References. (1) Abdo et al. 2010k; (2) Abdo et al. 2010l; (3) Abdo et al. 2010b; (4) Grondin et al. 2011; (5) Aharonian et al. 2006; (6) Aharonian et al. 2008; (7) Bird et al. 2010; (8) Castro & Slane 2010; (9) Cusumano et al. 2010; (10) Ghirlanda et al. 2010; (11) Mahony et al. 2010; (12) Mirabal 2009; (13) Mirabal et al. 2010; (14) Stephen et al. 2010; (15) Voges et al. 1994, 1999, 2000a, 2000b; (16) Watson et al. 2008.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

1FGL unassociated sources. These have also been added to the table. Where possible, names have been adjusted to be consistent with the NASA/IPAC Extragalactic Database[73] nomenclature.

4.2. Radio Searches for Pulsars

Of the 56 γ-ray emitting pulsars identified in 1FGL, 32 were detected by folding the γ-ray data using timing solutions from observations of known radio pulsars. These ephemerides have been collected by a global consortium of radio astronomers who dedicate a portion of their time toward this effort (Smith et al. 2008). The 32 pulsars (23 young and nine MSPs) had all been discovered in the radio band prior to their detection by the LAT. Since the release of the 1FGL catalog, 12 more of the 1FGL sources have been found to have γ-ray pulsations by using ephemerides of known radio pulsars. While 8 of the 12 were associated in 1FGL with a pulsar or an SNR/PWN, four were listed as unassociated sources in the catalog (two young pulsars and two MSPs). These new associations are listed in Table 7.

In addition to folding data using the properties of known radio pulsars, a promising technique for identifying unassociated sources is searching for previously unknown radio pulsars that might be powering the γ-ray emission. This technique was used on many of the EGRET unidentified sources (Champion et al. 2005; Crawford et al. 2006; Keith et al. 2008, for example) with modest success, because the error boxes were many times larger than a typical radio telescope beam. With the LAT, the unassociated source localizations are a much better match to radio telescope beam widths and generally each can be searched in a single pointing.

Thus far, over 450 unassociated LAT sources, mostly at high Galactic latitudes, have been searched by the *Fermi* Pulsar Search Consortium (PSC; Ransom et al. 2011) at 350, 820, or 1400 MHz. The target lists for these searches were selected from the LAT unassociated sources, with preference for those that displayed low variability and a spectrum consistent with

an exponential cutoff in the few GeV range, as seen in the identified γ-ray pulsar population (Abdo et al. 2010h). This program has resulted in the discovery of 32 previously unknown radio pulsars (one young pulsar and 31 MSPs; Ray & Parkinson 2011; Keith et al. 2011; Cognard et al. 2011; P. Bangale et al. 2012, in preparation) that are included in Table 7. Of these 32 new pulsars, 14 also show pulsations in γ rays. There is no obvious correlation between the γ-ray and radio fluxes of these pulsars, so additional discoveries can be expected as fainter unassociated 1FGL sources are searched.

Searches by the PSC continue. In the Galactic plane, high dispersion measures and sky temperatures demand higher frequency observations with concomitantly smaller beam sizes. Young, energetic pulsars can be very faint in the radio (Camilo et al. 2002a, 2002b). Nevertheless, we expect that deep observations will continue to turn up more discoveries of radio pulsars in unassociated 1FGL sources in the near future.

To summarize, radio observations of sources in the 1FGL unassociated sample have produced 36 new pulsar associations. Eighteen of these pulsars are considered firm identifications due to the detection of pulsations in the LAT data.

4.3. X-Ray Observations of Unassociated Source Fields

To look for additional possible counterparts we cross-correlated the list of unassociated 1FGL sources with existing X-ray source catalogs. We stress that the resulting compilation has no claim of completeness since the match with cataloged X-ray sources depends on the serendipitous sky coverage provided by the X-ray observations and the integration time of the observation. While it is possible that candidate X-ray counterparts to the LAT unassociated sources may be singled out on the basis of, e.g., their brightness and/or spectral properties, most will be recognized only through a coordinated multiwavelength identification approach (which is beyond the scope of this paper).

To begin, we considered the 2XMM source catalog derived from pointed *XMM-Newton* observations (Watson et al. 2008).

---

[73] http://ned.ipac.caltech.edu/

The fourth incremental release of the catalog (2XMMi) includes 191,870 unique X-ray sources extracted from all *XMM*-Newton observations that were public as of 2008 May 1, i.e., performed through the end of 2007 April. We cross-correlated the LAT source lists with the 2XMMi catalog, using a cross-correlation radius equal to the semimajor axis of the 95% confidence ellipse of the LAT source, and found that 40 of the 1FGL unassociated fields contained 2XMMi detections. Of these 40, four had been found to be associated with AGNs by the radio follow-up observations (Section 4.1).

By looking at the *XMM-Newton* observation log (up to 2011 February 27) we can estimate the potential increase in the number of matches that could occur if we were to use the longer observational database. For the cross-correlation we used the LAT 95% semimajor axis and compared against a radius equal to the sum in quadrature of the EPIC camera radius (15′) and the r95 positional uncertainty for the X-ray source. Here we found another 17 fields of unassociated LAT sources that have been observed by *XMM-Newton*, either serendipitously or pointed.

In addition, we cross-correlated with the *ROSAT* All Sky Survey Bright and Faint Source Catalogs (Voges et al. 1999, 2000b) and with the *ROSAT* catalog of pointed Position-Sensitive Proportional Counter (Voges et al. 1994) and High Resolution Imager (Voges et al. 2000a) observations. These found 101 unassociated source fields with *ROSAT* counterparts, 15 of which were also found in the 2XMMi correlation. These results are summarized in Table 7 and show that a preliminary X-ray screening provides potential X-ray counterparts for about 20% of all the *Fermi*-LAT unassociated sources. These possible X-ray counterparts are obviously prime targets for multiwavelength follow-up observations.

We also compared the 1FGL unassociated source list with recently published catalogs of hard X-ray/soft γ-ray sources. These are the Palermo *Swift*-BAT Hard X-ray Catalog (Cusumano et al. 2010), which is a compilation of 754 sources detected by the BAT instrument in its first 39 months of operations, and the 4th IBIS/SGRI Soft γ-ray Survey Catalog (Bird et al. 2010), which includes a total of 723 sources. Both catalogs contain flux and spectral information and provide likely source identification/classification. We cross-correlated the *Swift*-BAT and IBIS/SGRI source catalogs using the nominal 1FGL position and 95% semimajor axis as before, and found 11 new associations and eight new X-ray detections (cases where the candidate X-ray counterpart is not a known γ-ray emitter). The 1FGL 95% positional error is larger than the positional errors on sources in both catalogs. These results are also included in Table 7, where we give the *Swift*-BAT and Integral-IBIS/SGRI source identifications and proposed catalog classifications.

We note that a preliminary cross-correlation of the LAT unassociated sources with the *ROSAT* sources has been performed by Stephen et al. (2010). However, they used only the *ROSAT* Bright Source Catalog as a reference and found, on statistical grounds, that 60 of the 77 correlated positions should be genuine associations. However, they provide likely associations for only 30 of the correlations, those with counterparts within 160″. Table 7 lists 26 of these correlated sources. Three are not listed because the counterpart source type is not a known γ-ray emitter. The final source is not listed because no counterpart name was provided.

An analysis of archival *Swift* data for 21 unassociated 1FGL source fields was carried out by Mirabal (2009). This investigation indicated X-ray detections for seven LAT unassociated sources based on positional correlation with *Swift*-XRT sources

and the likelihood the source is a member of a γ-ray emitting class. Three of these are unique to this investigation and have been included as X-ray detections. In addition, Mirabal et al. (2010) proposed nine possible associations for unassociated LAT sources at $|b| > 25°$ in the 3000 deg^2 "overlap region," a region covered by various radio surveys and by the Sloan Digital Sky Survey. Associations and detections from both these investigations are included in Table 7.

The unassociated source 1FGL J1958.9+3459 appears to be nearly coincident with the HMXB Cygnus X-1, which was recently reported as an *AGILE* source (Sabatini et al. 2010). While this source meets the criteria to claim a positional association, there is no clear evidence that the source detected by the LAT is Cyg X-1. In addition, the source 1FGL J1045.2−5942 is positionally coincident with the luminous blue variable (LBV) star, Eta Carinae (η Car). While X-ray observations of η Car show a 5.54 year periodicity, the γ-ray flux remained constant during the most recent X-ray minimum in 2008 December–2009 January. However, due to its unusual γ-ray spectrum this 1FGL source is still believed to be associated with η Car (Abdo et al. 2010i).

To date, X-ray observations have led to positional associations with 10 AGNs, seven HMXBs, one SNR, and the LBV Eta Carinae. An additional 110 sources have X-ray detections that are excellent targets for follow-up multiwavelength observations. These associations can be found in Table 7.

4.4. TeV Observations of Unassociated Sources

Fermi-LAT spectra have been shown to be good predictors of TeV emission, with 55 1FGL sources having very high energy (VHE) counterparts (Abdo et al. 2009c, 2010j). The energy range from ∼50 GeV to ∼300 GeV is the only range where the LAT data directly overlap with other instruments.

The LAT team has provided recommendations for follow-up observations of a number of hard-spectrum sources—including unassociated hard-spectrum sources—that may have VHE counterparts. Coordinated follow-up observations in the TeV regime have been useful in identifying LAT-detected AGNs (see, e.g., Abdo et al. 2009c; Mariotti & MAGIC Collaboration 2010). In addition, LAT sources have been identified as SNRs by comparing the extension in the LAT data to the VHE emission by using the same procedure as was used for W51C (Abdo et al. 2009a), W44 (Abdo et al. 2010d), and IC443 (Abdo et al. 2010e). This search has yielded two more identified SNRs, W28 (Abdo et al. 2010k) and W49B (Abdo et al. 2010l). These associations can be found in Table 7.

We cross-checked the 1FGL unassociated source list with the list of TeV sources from TeVCat[74] and current publications. We consider a source to be coincident with a LAT source if its extension overlaps with the 95% confidence ellipse of the LAT source. We find nine TeV sources that are coincident with 1FGL unassociated sources (Table 8). Note that the 1FGL association process did not assign an association to a coincident TeV source if that TeV source had no identification in another waveband. Pismis 22 and W43 are possible (but not confirmed) associations with the TeV source.

One source of note is 1FGL J1702.4−4147, which lies on the emission "tail" of the elongated VHE source, HESS J1702−402. The nearby pulsar PSR J1702−4128 lies at the edge of the TeV γ-ray emission and would provide enough spin-down energy loss to produce the observed VHE emission via an extremely

---

[74] http://tevcat.uchicago.edu

Table 8
Candidate VHE Counterparts

1FGL Name	VHE Source Name	VHE Association	Reference
J0648.8+1516	VER J0648+152	Unidentified	Ong et al. (2010)
J1503.4−5805	HESS J1503−582	Unidentified	Renaud et al. (2008); Tibolla et al. (2009)
J1614.7−5138c	HESS J1614−518	*Pismis 22*	Aharonian et al. (2006)
J1702.4−4147	HESS J1702−402	PWN of PSR J1702−4128	Aharonian et al. (2008)
J1707.9−4110c	HESS J1708−410	Unidentified	Aharonian et al. (2006, 2008)
J1839.1−0543	HESS J1841−055	Multiple sources	Aharonian et al. (2008)
J1837.5−0659c	HESS J1837−069	Unidentified	Aharonian et al. (2006)
J1844.3−0309	HESS J1843−033	Unidentified	Hoppe (2008)
J1848.1−0145c	HESS J1848−018	*W43*	Tibolla et al. (2009)

Notes. Candidate VHE counterparts and their associations. Uncertain associations are in italics.

asymmetric PWN (Aharonian et al. 2008). Hence the pulsar is considered to be a likely counterpart to the LAT source. To date there has not been a high significance detection of pulsations in the LAT data. An additional interesting source, 1FGL J1839.1−0543, is positionally coincident with HESS J1841−055, one of the most extended (1 deg in diameter) HESS unidentified sources. Because of the high density of potential counterpart sources in this low-latitude region, there are multiple possible associations for the VHE source (two SNRs, three high spin-down PSRs, one XRB; Aharonian et al. 2008). Given its high TeV γ-ray flux, it is considered a good candidate for LAT detection (Tibolla 2009).

4.5. Association of LAT Sources Using Only LAT Data

A small number of sources have been associated or identified since the release of the 1FGL catalog by using LAT data alone. Of the 56 pulsars listed in 1FGL, 24 were discovered using blind frequency searches (Abdo et al. 2009d) for γ-ray pulsations from the bright unassociated sources. These are typically young pulsars, for which the solid angle of the radio beam is likely to be much smaller than the γ-ray one (Abdo et al. 2010h). As a result of this geometry, many unassociated sources are likely to be young, radio-quiet pulsars that will never be found in radio searches. Since the release of the 1FGL catalog, another blind search pulsar, PSR J0734−1559, has been discovered in an unassociated LAT source. This association has been included in Table 7.

One new HMXB has also been detected in the LAT data (Corbet et al. 2011), though in 1FGL the source was associated with the SNR G284.3−01.8. This is the first of its type to be discovered in γ rays. These new associations are also included in Table 7.

5. DISCUSSION

The follow-up multiwavelength association efforts discussed in Section 4 have resulted in 177 new extragalactic source associations (all AGNs), and 52 new Galactic associations (one source has both a Galactic and an extragalactic association). When we compare these new associations with the expected 1FGL source distributions discussed in Section 2.4, the estimated numbers of sources that have not yet been associated are reduced to 182 for extragalactic sources and 219 for Galactic sources.

To test the two classification algorithms and to estimate the efficiency for identifying the different source classes, we compare the results to the new source associations in Section 4, first looking at individual methods, and then the combined

results. In addition, we consider how the results match the log N–log S analysis.

5.1. Validating the Classification Results from the Separate Methods

We separately compared the results of the Classification Tree analysis and the LR analysis to the new source associations in Section 4. Altogether, the new source associations include 177 new AGN associations as well as 37 new pulsar associations that are divided into two categories: 20 objects for which pulsations have been detected in γ rays (which we will refer to as "new pulsar detections") and 17 objects for which pulsations have been detected only in the radio (which we will refer to as "new pulsar candidates"). We will not use new associations from other source types (HMXBs, PWNe, SNRs) for this validation.

For AGNs, we find that 126 sources are correctly classified as AGN candidates by the CT analysis (efficiency: 71%), 11 were classified as pulsar candidates (false negative: 6%), while the remaining 40 sources were considered still unclassified (23%). The same comparison for the LR analysis gave 142 sources correctly classified as AGN candidates (efficiency: 80%), seven sources classified incorrectly as pulsar candidates (false negative: 4%), while the other 28 sources remained unclassified (16% of the sample).

For pulsars, we noticed a very different performance between new pulsar detections and new pulsar candidates. For the 20 sources detected as pulsars by the LAT, the CT analysis correctly classifies 14 pulsars (efficiency: 70%), misclassifies one source (5%), and leaves the remaining sources unclassified (25%). The LR analysis correctly classifies 11 pulsars (efficiency: 55%), misclassifies one source (5%), and leaves the remaining sources unclassified (40%). On the other hand, for the new radio pulsar candidates, the CT analysis correctly classifies only three objects as pulsars (efficiency: 18%), misclassifies eight objects as AGNs (false negative: 47%), and leave the six remaining objects as still unassociated (35% of the new pulsar candidates). The LR analysis correctly classifies four pulsar candidates (efficiency: 23.5%), misclassifies four sources (23.5%), and leaves the remaining nine sources unclassified (53%).

These results are interesting, as the definition of the pulsar fiducial threshold in the LR analysis appeared that it might overestimate the pulsar candidates. However, the LR actually has a somewhat poorer success rate for finding new pulsars and pulsar candidates than the Classification Tree analysis. Looking more closely at the 1FGL properties of the misclassified pulsar candidates, we found that 12 of the 17 new pulsar candidates have only upper limits for the 300 MeV–1 GeV band. In contrast, 80% of the new pulsar detections were significantly
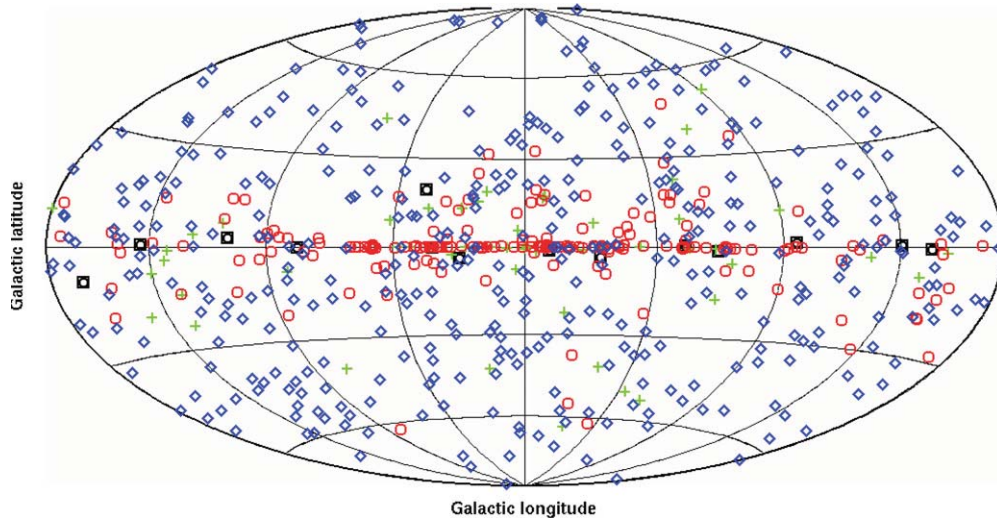
Figure 11. Spatial distribution of the combined classification sample, in Galactic coordinates. Sources are classified as AGN candidates (blue diamond), pulsar candidates (red circles), unclassified (green crosses), or in conflict (black squares).

(A color version of this figure is available in the online journal.)

detected in this portion of the LAT spectrum. This difference in characteristics for the two pulsar groups may indicate the need for additional criteria when selecting sources for follow-up observations.

5.2. Validating the Results from the Combined Classifications

We can also compare the new associations to the combined classifications defined in Section 3.4. Of the 214 newly associated AGNs and pulsars from Section 4, 171 sources (151 new AGNs, 16 new pulsar detections, and four new pulsar candidates) match the classification given by the combined analysis, and 26 sources (15 new AGNs, one new pulsar detection, and 10 new pulsar candidates) are in direct conflict with the classification source type. This gives an efficiency of 85% for AGN classification and 80% for classification of new pulsar detections, but only 59% for new pulsar candidates. Seventeen of the newly associated sources are unclassified by either method, and only one source has conflicting source classification. The one conflicting source turns out to be a new pulsar candidate that also has an AGN association, suggesting the LAT source could be the sum of these two objects. The overall efficiency for this combined sample is ∼80%, comparable to the value we were seeking when we set the fiducial values for the two methods. The combined sample has a false negative rate of ∼12%.

The spatial distributions of the newly classified sources give us the opportunity to cross check our results. Figure 11 shows the spatial distribution of all the sources. Note that both the AGN and pulsar distributions are as expected, even though we have not used the Galactic latitude as an input to either classification method. The pulsar candidates are mainly distributed along the Galactic plane, with a few high-latitude exceptions that suggest additional nearby MSPs, while the AGN candidates are nearly isotropically distributed on the sky.

From this we can conclude that using only the γ-ray properties of the *Fermi*-LAT sources, and the firm associations of the 1FGL, we were able to develop a prediction for AGN and pulsar classification that nearly matches our expectations (i.e., pulsar candidates are not variable, have a curved spectrum, and are mainly distributed along the Galactic plane, while AGN

candidates are mostly extragalactic, variable sources). In all, the efficiency of the combined classification methods at classifying new AGNs is high, with a low rate of false negatives, while the efficiency for new pulsar candidates is much lower than expected.

AGNs and pulsars are not the only γ-ray source classes known or expected, but the less-populous source types are hard to classify using the techniques described here because the training samples are too small in 1FGL. With time, those training samples will likely grow, and we may be able to extend this analysis to additional classifications.

5.3. Comparison to the log N–log S Predictions

In addition, we can check to see how the classification results compare to the predictions made by the $\log N$–$\log S$ analysis in Section 2.4. For this comparison, we consider a pulsar classification to be indicative of a Galactic source, and an AGN classification to be indicative of an extragalactic source.

Since 229 of the unassociated sources now have associations, we will consider only those 401 that remain unassociated. Thirty-three of these are unclassified by either technique and 13 have conflicting classifications. The number sources for low latitudes and high latitudes for the remaining 355 are shown in Table 2. At high latitudes, the observed numbers of both Galactic and extragalactic sources are consistent with the numbers expected from the $\log N$–$\log S$ analysis. In contrast, at low latitudes, the number of Galactic sources is about one-third lower than expected and the number of extragalactic sources is higher than expected. It is clear that the group of sources that is hardest to associate is those of Galactic origin at low latitudes, likely due to the presence of a population of spurious sources in that region in the 1FGL catalog.

Figure 12 (left panel) shows the latitude distribution of the classified unassociated sources based on classification type. If we combine the AGN candidate population with the 1FGL sources that already have AGN associations (Figure 12, right panel), we find that the shape of the AGN distribution matches reasonably with that predicted by the model, though there is still an excess at low Galactic latitudes.
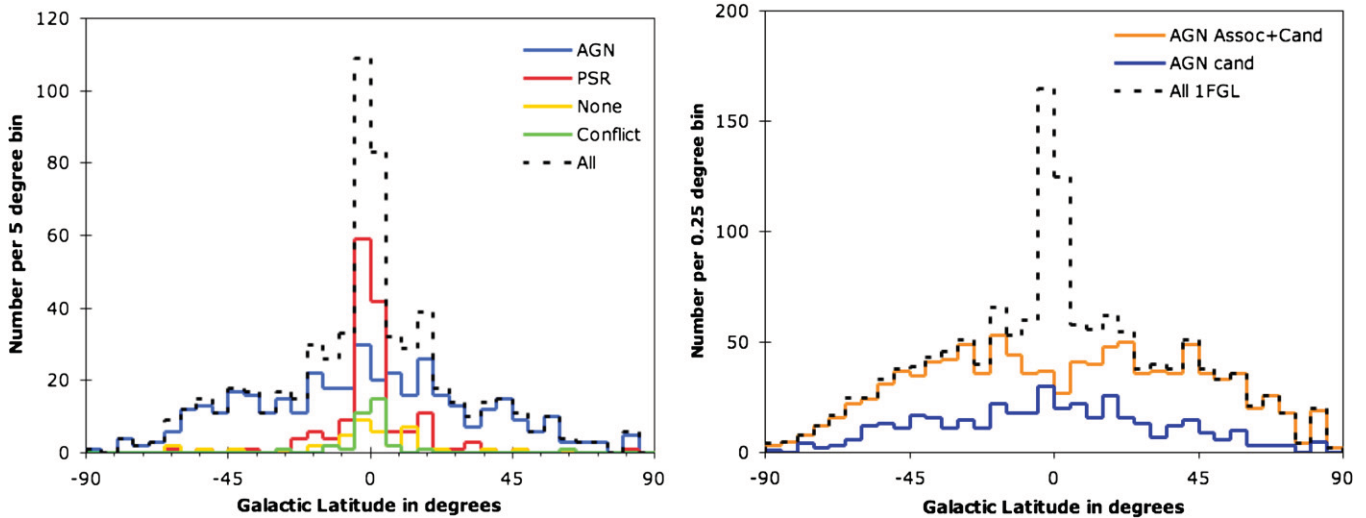
Figure 12. Left: distribution of classified sources binned by Galactic latitude, with AGNs in blue, pulsars in red, unclassified sources in green, and sources with conflicting classification in yellow. The dashed line is the total distribution. Right: distribution of AGN candidate binned by Galactic latitude. The orange line is the sum of the 1FGL AGN associations plus the sources classified as AGN candidates (blue line). The dashed line is the distribution for all 1FGL sources.

(A color version of this figure is available in the online journal.)
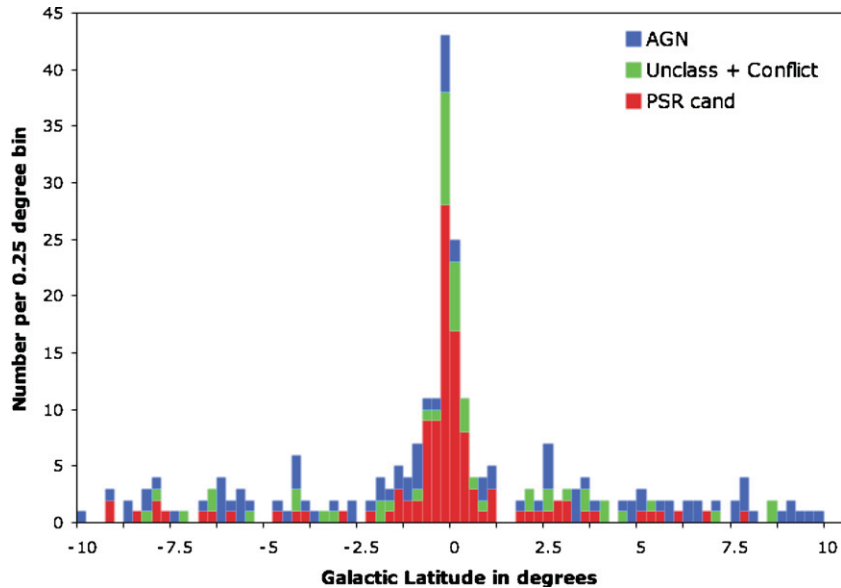
Figure 13. The 1FGL unassociated sources in the central few degrees of the Galaxy can be mostly separated into those classified as pulsars (red) and those that have conflicting classifications or were unable to be classified (green). The few remaining sources were classified as AGN (blue).

(A color version of this figure is available in the online journal.)

5.4. Unassociated Sources at Low Galactic Latitudes

It is the unassociated sources in the central $20°$ of Galactic latitude ($|b| < 10°$) that may hold clues to the content of the narrow Galactic ridge population at $|b| < 0°.5$, $|l| < 60°$ discussed in Section 2.1. To investigate these sources, we separate the pulsar candidates from the other types of sources (Figure 13) and consider the distribution. Where the full set of unassociated sources appears to indicate an unreasonably narrow scale height of ∼50 pc for the population, the latitude distribution of pulsar candidates is somewhat closer to expectations, implying a scale height of ∼85 pc. This value is one-third the scale height of LAT-detected γ-ray pulsars (Abdo et al. 2010h).

For γ-ray sources, this population scale height suggests instead that Population I objects such as SNRs, with a scale height of ∼100 pc (Lamb & Macomb 1997), are likely

significant contributors to the 1FGL sources. In the 1FGL catalog, 44 sources were associated with or identified as SNRs, and this paper associates six more (Table 7), giving a total of 50 SNR associations. Considering that only 63 1FGL sources were associated/identified with pulsars, it is clear that both source types are significant contributors. The classification method used here does not consistently label SNRs as pulsar candidates. Of the six new SNR associations, five are classified as pulsar candidates and the sixth is classified as an AGN. In the future it may be useful to consider the SNR source class separately as an input to such classification analyses.

This central portion of the Galaxy is also the region that has most of the sources that are either unclassified or have conflicting classifications. Of the 257 unassociated sources in this region, 22 have no classification and 29 have conflicting classifications between the two methods. Of these 51 unclassified or

18

conflicting-classification sources, eight have new associations. The source types are varied; two MSPs, one young pulsar, three HMXBs, and two AGNs. It is clear that not all of these sources can be spurious, however it is unlikely that the remaining 43 are all real detections.

5.5. Informing Future Follow-up Observations

The results of the classification analyses demonstrate that source properties measured with the *Fermi*-LAT can provide important guidance on what types of follow-up observations are likely to be fruitful for many of these unassociated sources. The emphasis in follow-up observations of LAT sources has been on radio imaging and timing observations for a large number of sources, as well as targeted X-ray observations for sources of interest (e.g., flaring sources or new radio pulsar candidates). In addition, there is an on-going program to observe all the bright, well-localized *Fermi*-LAT unassociated sources with *Swift*[75] that may add important new insights into these sources as a group.

In Table 4, the last four columns show what follow-up observations are recommended in several wavebands. Obviously, sources classified as likely blazars would benefit from radio searches for flat-spectrum sources within the LAT error ellipse. Low-frequency radio timing is recommended for the likely pulsars. X-ray observations of likely pulsars can give timing observers seed locations at which to search for pulsations in both radio and LAT data (Caraveo 2009). Still-unassociated sources that may benefit from such observations have been flagged in the 1FGL unassociated source list with the appropriate observation type. These are suggestions; it is highly likely that some of the sources are misclassified. Also, a number of the follow-up observations discussed here have yielded no new associations for some of the observed unassociated sources.

We strongly recommend additional joint analyses of LAT and ground-based VHE γ-ray data for very low latitude ($|b| < 0°5$) *Fermi* sources. Together, these may give insight into whether or not a population of SNRs can account for a significant number of the 1FGL unassociated sources along the Galactic ridge. Sources for which this type of analysis is recommended are indicated in the 1FGL unassociated source list (Table 4).

5.6. Remaining BSL Unassociated Sources

Follow-up observations like those discussed in the previous section have made a significant impact, increasing the associated fraction for the 1FGL catalog from ~56.5% to ~71% in only a little more than one year. But we can look farther back to the *Fermi*-LAT Bright Source List (BSL; Abdo et al. 2009e), the list of 205 high significance ($>10\sigma$) LAT sources detected in the first three months of the *Fermi* mission. Of the 37 sources listed as unassociated in the BSL, 10 now have pulsar identifications or associations, while eight have new AGN associations. In addition, seven have been associated with other γ-ray source types such as SNRs or HMXBs. These associations bring the BSL association rate up to 94%.

Of the remaining twelve BSL unassociated sources, five lie in the Galactic ridge and should be considered with caution. Here, we look more closely at these 12 sources.

1. *1FGL J0910.4−5055 (0FGL J0910.2−5044).* This Galactic plane transient ($l, b = 271°7, −1°96$) flared once in 2008 October, an event lasting 1–2 days with peak flux

($E > 100$ MeV) of 1×10^{-6} photons cm^{-2} s^{-1} (Cheung et al. 2008). (1FGL reported a peak flux for this source of 1.97×10^{-7} photons cm^{-2} s^{-1}, but that was averaged over one month.) Recent figure-of-merit analysis by Murphy et al. (2010) has given this source a >80% probability of being associated with the radio source AT20G J091058-504807. Our analysis does not find such an association, though Mirabal (2009) does show an association with the likely blazar *Swift* J0910.9−5048.

2. *1FGL J1311.7−3429 (0FGL J1311.9−3419).* Besides being associated with 3EG J1314−3431, this very bright high-Galactic latitude ($l, b = 307°7, 28°2$) source is not variable and has a spectrum with a high-energy cutoff very similar to a pulsar. To date, searches for both γ-ray and radio pulsations from this source have been unsuccessful (Ransom et al. 2011).

3. *1FGL J1536.5−4949 (0FGL J1536.7−4947).* With no significant variability, this persistently bright mid-latitude ($l, b = 328°2, 4°8$) source is the only unassociated BSL source to have conflicting classifications. This source has a moderately curved spectrum with a high-energy cutoff.

4. *1FGL J1620.8−4928c (0FGL J1622.4−4945).* This moderately bright source in the Galactic ridge ($l, b = 333°9, 0°4$) is spatially coincident with the *AGILE* detection 1AGL J1624-4946 and has a spectrum with a sharp spectral break at ~3 GeV. There is no evidence at this time for pulsations from this source in either γ rays or radio (Ransom et al. 2011).

5. *1FGL J1653.6−0158 (0FGL J1653.4−0200).* Another non-varying source with a pulsar-like spectrum, this high-latitude ($l, b = 16°6, 24°9$) source is associated with 3EG J1652−0223. Both classification methods call this source a pulsar candidate.

6. *1FGL J1740.3−3053c (0FGL J1741.4−3046).* The position for this non-varying c-source in the Galactic ridge ($l, b = 357°7, −0°1$) moved far enough between the two publications that the two detections are not formally associated. However, we recall that for pulsars in the plane, the 1FGL error ellipse appears to underreport the systematic error. As the BSL source lies just outside the 1FGL 95% confidence contour, we consider the two detections to be related.

7. *1FGL J1839.1−0543c (0FGL J1839.0−0549).* This bright source in the Galactic ridge ($l, b = 26°4, 0°1$) has a highly curved spectrum and does not vary. Both classification methods call this source a pulsar candidate.

8. *1FGL J1842.9−0359c (0FGL J1844.1−0335).* This Galactic ridge source ($l, b = 28°4, 0°1$) had a source significance of 10σ in the first three months, but after 11 months that value has increased only slightly to 10.9σ (unless variable, a source should have twice the significance after nearly four times the livetime). Since 1FGL found this source to be non-varying, it is unlikely that the low flux in 1FGL is due to variability effects. Instead, it appears that the longer data set has separated the BSL source into multiple components, leaving the coincident 1FGL source at a lower-than-expected significance.

9. *1FGL J1848.1−0145c (0FGL J1848.6−0138).* Another source in the Galactic ridge ($l, b = 31°0, −0°1$), its spectrum is consistent with a power-law, and the source may be related to a TeV source (see Table 8).

10. *1FGL J2027.6+3335 (0FGL J2027.5+3334).* A bright, mildly variable source that lies near the Galactic plane

$(l, b = 73°.3, −2°.9)$, this source is associated with the EGRET source 3EG J2027+3429. In the 1FGL catalog, this source flux peaked at $3.4 × 10^{-7}$ photons cm^{-2} s^{-1} with a significance $>10\sigma$ in a single month. The spectrum has a large curvature index, indicating that it is not a simple power law. Although the variability seems to indicate a possible AGN, both classification methods consider it a likely pulsar.

11. *1FGL J2111.3+4607 (0FGL J2110.8+4608).* While this source near the Galactic plane $(l, b = 88°.3, −1°.4)$ is highly significant, it is flagged in the 1FGL catalog as having the flux measurement that is sensitive to changes in the diffuse model. Even so, the spectrum is moderately pulsar-like with a high-energy cutoff and there is no hint of variability. Both classification techniques consider this source a pulsar candidate.

12. *1FGL J2339.7−0531 (0FGL J2339.8−0530).* This very hard source (spectral index = 1.99) lies at high Galactic latitude $(l, b = 81°.4, −62°.5)$. While its 1FGL five-band spectrum suggests a blazar, neither classification method was able to classify this source.

While five of these are c-sources, only two (1FGL J1842.9−0359c and 1FGL J1848.1−0145c) appear to be questionable, though the possibility of a TeV component for the latter should be investigated. The two variable sources seem likely to be AGNs. The γ-ray characteristics of the majority of the other unassociated BSL sources imply that these sources are bright, steady, and have curved spectra with high-energy cutoffs. With the exception of 1FGL J0910.4−5055, all these sources were included in the searches for radio pulsations, the same searches that have resulted in the discovery of 10 new MSPs in unassociated BSL sources. Blind searches for pulsations in the γ-ray data have also been performed on these sources, with no success.

One source of interest is 1FGL 2017.3+0603, a BSL source that now has detected pulsations in the LAT data (Cognard et al. 2011) as well as an AGN association with a high figure of merit (0.923) in 1LAC (Abdo et al. 2010b). This source will require additional analysis to determine whether the LAT flux is due solely to the pulsar or is a combination of both counterparts.

5.7. Comparing with EGRET Unassociated Sources

Although the present paper is focused on the *Fermi*-LAT unassociated sources, some insight about these sources may be found from the all-sky survey with EGRET on the *Compton Gamma Ray Observatory*. Using the LAT results, with higher sensitivity and better source locations, as a reference, we re-examine the sources of two EGRET catalogs: the 3EG catalog (Hartman et al. 1999) and the EGR catalog (Casandjian & Grenier 2008). The two catalogs were based largely on the same data, while they used different models of the diffuse γ radiation that forms a background for all sources. In addition, noting those EGRET unassociated sources that remain unassociated in the 1FGL catalog offers the opportunity to recognize sources that remain interesting on a timescale of decades.

5.7.1. Experience with the EGRET Catalogs

A comparison of the two EGRET catalogs with each other and with the LAT 1FGL catalog yields several observations and tentative conclusions, illustrated by specific examples. Because more detailed information is available about the 3EG sources, many of these results emphasize that catalog, although similar considerations likely apply to the EGR analysis.

1. The statistical error contours produced for EGRET underestimated the full uncertainty in the source localization. The Crab, Vela, Geminga, and PSR J1709−4429 pulsars, which are positively identified by γ-ray pulsations in both EGRET and *Fermi*-LAT data, have positions outside the formal error contours, even at the 99% level, as noted by Hartman et al. (1999). For 3EG, but not for EGR, the CTA 1 and LSI +61° 303 sources, now firmly identified by LAT, also lie outside the 95% error contours. For this reason, we expand the list of plausible associations to include all those LAT sources whose position falls within the 99% error contours. As noted by Hartman et al. (1999), however, the EGRET source localizations were better at higher Galactic latitudes. Even bright AGNs such as 3C279 were typically found within the error contours.

2. The circular fit to the EGRET 95% error contour was a poor approximation in many cases. In addition to the 107 matches between 1FGL and 3EG reported based on the automated comparison using this circular fit, there are 21 more 1FGL sources found within the 95% error contours of the detailed 3EG uncertainty contour maps (Hartman et al. 1999). Further, 19 1FGL sources are located between the 3EG 95% and 99% contours, making a total of 149 candidate associations or 153 including the four bright pulsars identified by timing. The 14 other 1FGL sources that lie just outside the 3EG 99% contours are not included in this analysis, although they remain potential association.

3. The adopted diffuse background model is important both close to and far from the Galactic plane. For example, EGR J0028+0457 is confirmed by the LAT as the MSP PSR J0030+0451 (Abdo et al. 2009f) at a Galactic latitude of $−57°$. There was no corresponding source in the 3EG catalog, though a sub-threshold excess was seen (D. J. Thompson 2010, private communication). Conversely, 3EG J1027−5817, at a Galactic latitude of $−1°$, is confirmed by the LAT as PSR J1028−5819 (Abdo et al. 2009g), while the nearest EGR source is nearly $1°$ away with a 95% error uncertainty of $0°.22$.

4. Variability is an important consideration even for sources not associated with blazars. The EGRET upper limit for radio galaxy NGC 1275, derived from Figure 3 of Hartman et al. (1999), lies nearly an order of magnitude below the LAT detection level (Kataoka et al. 2010). 3EG J0516+2320 was a bright solar flare from 1991 June (Kanbach et al. 1993). The 1FGL catalog contains no solar flares. 3EG J1837−0423 (Tavani et al. 1997) was one of the brightest sources in the γ-ray sky in 1995 June, at a flux ($E>$ 100 MeV) of $(3.1 ± 0.6) × 10^{-6}$ photons cm^{-2} s^{-1}. No 1FGL source is seen consistent in position with this source, even though the LAT would have detected a source nearly 2 orders of magnitude fainter (see Figure 19 of Abdo et al. 2010a).

Some of these lessons from the EGRET era have already been applied to the 1FGL catalog construction but are worth reiterating in any discussion of unassociated sources.

1. Unlike EGRET, the error contours for 1FGL include a systematic component of 10% added to the statistical uncertainties. This component was derived, however, from high-latitude AGN comparisons. The EGRET experience suggests that the situation will be more difficult at lower Galactic latitudes. The low-latitude LAT source associated with LS5039 based on periodicity, for example, has a

measured position that lies at the 95% uncertainty contour even after adding a 20% systematic component (Abdo et al. 2009h). This is an additional indication that the LAT positional uncertainties for sources in the Galactic plane are affected by the systematics discussed earlier.

2. Even without any additional systematic uncertainty, it should be remembered that at least 5%, or more than 70, of the 1FGL sources probably have true counterparts that fall outside the 95% contours.

3. Just as the EGRET catalog used "C" for potentially confused regions and "em" for possibly extended or multiple sources, the 1FGL catalog identifies sources with a "c" in the name or a numerical flag to indicate possible uncertainties due to the analysis procedure or the diffuse model. The LAT has confirmed ∼67% of EGRET sources that do not carry either the "C" or "em" flags, but has only confirmed ∼38% of the flagged sources. This experience with EGRET certainly suggests that such flags should be taken seriously.

5.7.2. 1FGL Unassociated Sources Remaining from the EGRET Era

Despite the gap of more than a decade between the EGRET and LAT observations, a sizable number of the unassociated sources seen by EGRET are associated with 1FGL sources that are unassociated with known source classes. Whether they are persistent or recurrent, such sources offer a high potential for multiwavelength studies. Table 7 includes 43 positional coincidences we have found between unassociated sources in the EGRET and 1FGL catalogs. As noted above, the predominance of 3EG sources results at least partly from the additional information available compared to the EGR catalog, which did not include confidence contour maps.

These unassociated sources are distributed widely across the sky. Only 10 of the 43 have "c" designations in 1FGL, and all of these lie close to the Galactic plane toward the inner Galaxy, where the Galactic diffuse emission is brightest and any deficiencies in the model of the diffuse emission would have the greatest effect on properties of the 1FGL sources. Although the EGRET localization uncertainties are large, the density of 1FGL sources away from the Galactic Plane is not so large that accidental coincidences are a significant problem. EGR J1642+3940 (1FGL J1642.5+3947) appears to be a special case. It appeared in the EGRET data only after the end of the 3EG data set, but it has been seen by the LAT. Casandjian & Grenier (2008) suggest an association with blazar 3C 345, although it might also be associated with Mkn 501 (Kataoka et al. 1999). Although it is shown as unassociated in the 1FGL catalog, recent LAT analysis also suggests one or more blazars contribute to this source (Schinzel et al. 2010).

6. CONCLUSIONS

As the *Fermi* mission matures, it is important to take a look at the successes of the early mission to help inform and improve the association and follow-up of new sources. The continued multiwavelength observations and ongoing statistical association efforts for the 1FGL unassociated sources have led to associations for 70% of the entire catalog. Since the release of the catalog, 45% of all the extragalactic sources expected from the log *N*–log *S* analysis have been associated. In addition, 47% of the expected high-latitude Galactic sources have been associated. Together, this gives associations for nearly 82% of all expected extragalactic and nearly 62% of expected high-latitude Galactic sources. However, there are associations for

only ∼ 38% of the expected sources of Galactic origin that lie at $|b| < 10°$.

The significant improvement in sensitivity of the *Fermi*-LAT relative to EGRET, the substantially improved positional errors, and the sky-survey viewing plan made possible by the large field of view of the LAT have generated an unprecedented data set and allowed the production of the deepest-ever catalog of the GeV sky. From that foundation, the astronomical community has worked in concert to discover new additions to known γ-ray source classes, as well as adding SNRs, PWNe, starburst galaxies, radio galaxies, HMXBs, globular clusters, and a treasure trove of MSPs to the list.

With that success as a backdrop, there are clearly lessons we have learned from the 1FGL catalog process and follow-up analyses, as well as from experience with previous missions.

1. As discussed in Abdo et al. (2010a), it is clear that there is room for improvement in the plane of the Galaxy and especially the ridge. This region contributes numerous questionable sources to the catalog. However, if we set that region aside, essentially half of the remaining 1FGL unassociated sources in other regions now have associations, giving an overall association rate of ∼70% for 1FGL. In contrast, the 3rd EGRET catalog had an association rate of only 38%.

2. Follow-up campaigns to associate or identify the LAT-detected sources have been extremely successful. A total of 178 new blazar candidate associations have been made, primarily by looking for radio candidates within the LAT error ellipses and then re-observing at additional frequencies to determine whether the source has the characteristic flat spectrum in the radio. Thirty-one new Galactic field MSPs have been discovered based on locations provided by the LAT for radio pulsations searches.

3. By using the distribution of detected sources, we can model the 1FGL γ-ray sky and predict how many of each general source type we expect in the catalog. After taking into account the associations made since the release of the 1FGL catalog, this analysis indicates that, as the mission continues, we might expect to find associations for at least ∼200 more AGNs (mostly at high Galactic latitudes) and ∼50 new pulsars (equally divided at high and low latitudes) among the unassociated 1FGL sources.

4. We have applied two analysis techniques to infer the likely classification for unassociated sources, based solely on their γ-ray properties. The γ-ray properties of sources, while not being sufficient on their own to determine source type, can provide important information regarding the parent source classes. Using the information provided by the LAT to inform the selection of γ-ray sources and wavelengths for follow-up studies can reduce the labor intensive nature of such observations and increase the likelihood of finding a viable association candidate. A preliminary assessment of this process shows a success rate of ∼80%.

Facility: Fermi

REFERENCES

Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009a, ApJ, 706, L1
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009b, ApJ, 700, 597
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009c, ApJ, 707, 1310
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009d, Science, 325, 840 (Blind Search Pulsars)
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009e, ApJS, 183, 46 (Bright Source List)
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009f, ApJ, 699, 1171 (PSR J0030+0451)
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009g, ApJ, 695, L72 (PSR J1028−5819)
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009h, ApJ, 706, L56 (LS5039)
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010a, ApJS, 188, 405 (1FGL Catalog)
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010b, ApJ, 715, 429
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010c, ApJ, 719, 1433
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010d, Science, 327, 1103
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010e, ApJ, 712, 459
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010f, ApJ, 720, 435
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010g, ApJ, 722, 520
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010h, ApJS, 187, 460
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010i, ApJ, 723, 649
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010j, ApJ, 711, 64
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010k, ApJ, 718, 348
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2010l, ApJ, 722, 1303
Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2011, Science, 331, 739
Ackermann, M., Ajello, M., Baldini, L., et al. 2010, ApJ, 721, 1383
Ackermann, M., Ajello, M., Baldini, L., et al. 2011, ApJ, 726, 35
Aharonian, F., Akhperjanian, A. G., Barres de Almeida, U., et al. 2008, A&A, 477, 353
Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2006, ApJ, 636, 777
Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, ApJ, 697, 1071 (LAT)
Bird, A. J., Bazzano, A., Bassani, L., et al. 2010, VizieR Online Data Catalog, 218, 60001
Bock, D., Large, M. I., & Sadler, E. M. 1999, AJ, 117, 1578
Breiman, L., Freidman, J. H., Olshen, R. A., & Stone, C. J. 1984, Classification and Regression Trees (new edition; Monterey, CA: Wadsworth and Brooks)
Browne, I. W. A., Wilkinson, P. N., Jackson, N. J. F., et al. 2003, MNRAS, 341, 13
Brun, R., & Rademakers, F. 1997, Nucl. Instrum. Meth., A389, 81
Camilo, F., Lorimer, D. R., Bhat, N. D. R., et al. 2002a, ApJ, 574, L71
Camilo, F., Stairs, I. H., Lorimer, D. R., et al. 2002b, ApJ, 571, L41
Caraveo, P. A. 2009, arXiv:0912.4857
Casandjian, J.-M., & Grenier, I. A. 2008, A&A, 489, 849
Castro, D., & Slane, P. 2010, ApJ, 717, 372
Champion, D. J., McLaughlin, M. A., & Lorimer, D. R. 2005, MNRAS, 364, 1011
Cheung, C. C., Reyes, L., Longo, F., & Iafrate, G. 2008, ATel, 1788, 1
Cognard, I., Guillemot, L., Johnson, T. J., et al. 2011, ApJ, 732, 47
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Corbet, R. H. D., Cheung, C. C., Kerr, M., et al. 2011, ATel, 3221, 1
Crawford, F., Roberts, M. S. E., Hessels, J. W. T., et al. 2006, ApJ, 652, 1499
Cusumano, G., La Parola, V., Segreto, A., et al. 2010, BAAS, 41, 715
Drlica-Wagner, A., & Charles, E. 2011, arXiv:1111.2352
Ghirlanda, G., Ghisellini, G., Tavecchio, F., & Foschini, L. 2010, MNRAS, 407, 791
Grondin, M.-H., Funk, S., Lemoine-Goumard, M., et al. 2011, ApJ, 738, 42
Hartman, R. C., Bertsch, D. L., Bloom, S. D., et al. 1999, ApJS, 123, 79
Healey, S. E., Fuhrmann, L., Taylor, G. B., Romani, R. W., & Readhead, A. C. S. 2009, AJ, 138, 1032
Healey, S. E., Romani, R. W., Taylor, G. B., et al. 2007, ApJS, 171, 61
Hocker, A., et al. 2007, TMVA—Toolkit for Multivariate Data Analysis, arXiv:physics/0703039
Hoppe, S. 2008, 30th ICRC Conf. (Mexico City), 2, 579
Hosmer, D., & Lemeshow, S. 2000, Applied Logistic Regression (2nd ed.; New York: Wiley)
Kanbach, G., Bertsch, D. L., Fichtel, C. E., et al. 1993, A&AS, 97, 349
Kataoka, J., Mattox, J. R., Quinn, J., et al. 1999, Astropart. Phys., 11, 149
Kataoka, J., Stawarz, Ł., Cheung, C. C., et al. 2010, ApJ, 715, 554
Keith, M. J., Johnston, S., Kramer, M., et al. 2008, MNRAS, 389, 1881
Keith, M. J., Johnston, S., Ray, P. S., et al. 2011, MNRAS, 414, 1292
Lamb, R. C., & Macomb, D. J. 1997, ApJ, 488, 872
Mahony, E. K., Sadler, E. M., Murphy, T., et al. 2010, ApJ, 718, 587
Mariotti, M., & MAGIC Collaboration. 2010, ATel, 2916, 1
Mattox, J. R., Bertsch, D. L., Chiang, J., et al. 1996, ApJ, 461, 396
Mirabal, N. 2009, arXiv:0908.1389
Mirabal, N., Nieto, D., & Pardo, S. 2010, arXiv:1007.2644
Murphy, T., Sadler, E. M., Ekers, R. D., et al. 2010, MNRAS, 402, 2403
Myers, S. T., Jackson, N. J., Browne, I. W. A., et al. 2003, MNRAS, 341, 1
Nolan, P. L., Abdo, A. A., Ackermann, M., et al. 2012, ApJS, 199, 31 (2FGL Catalog)
Ong, R. A., VERITAS Collaboration, Paneque, D., & Fermi Large Area Telescope. 2010, ATel, 2486, 1
Perlich, C., Provost, F., & Simonoff, J. S. 2003, J. Mach. Learn. Res., 4, 211
Ransom, S. M., Ray, P. S., Camilo, F., et al. 2011, ApJ, 727, L16
Ray, P. S., & Parkinson, P. M. S. 2011, in High-Energy Emission from Pulsars and Their Systems, ed. D. F. Torres & N. Rea (Berlin: Springer), 37
Renaud, M., Goret, P., & Chaves, R. C. G. 2008, in AIP Conf. Proc. 1085, High-Energy Gamma-ray Astronomy: Proc. 4th International Meeting on High Energy Gamma-Ray Astronomy, ed. F. A. Aharonian, W. Hofmann, & F. Rieger (Melville, NY: AIP), 281
Sabatini, S., Striani, E., Verrecchia, F., et al. 2010, ATel, 2715, 1
Schinzel, F. K., Lobanov, A. P., Jorstad, S. G., et al. 2010, in Proc. Fermi Meets Jansky—AGN in Radio and Gamma-Rays, ed. T. Savolainen et al. (Bonn: MPIfR), arXiv:1012.2820
Smith, D. A., Guillemot, L., Camilo, F., et al. 2008, A&A, 492, 923
Stephen, J. B., Bassani, L., Landi, R., et al. 2010, MNRAS, 408, 422
Swanenburg, B. N., Bennett, K., Bignami, G. F., et al. 1981, ApJ, 243, L69
Tavani, M., Bulgarelli, A., Vittorini, V., et al. 2011, Science, 331, 736
Tavani, M., Mukherjee, R., Mattox, J. R., et al. 1997, in AIP Conf. Proc. 410, Proc. Fourth Compton Symposium, ed. C. D. Dermer, M. S. Strickman, & J. D. Kurfess (Melville, NY: AIP), 1253
Thompson, D. J., Bertsch, D. L., Fichtel, C. E., et al. 1993, ApJS, 86, 629
Tibolla, O. 2009, in AIP Conf. Proc. 1112, Science with the New Generation of High Energy Gamma-ray Experiments: Proc. 6th Edition: Bridging the Gap Between GeV and TeV, ed. D. Bastieri & R. Rando (Melville, NY: AIP), 211
Tibolla, O., Chaves, R. C. G., Domainko, W., et al. 2009, arXiv:0912.3811
Vandenbroucke, J., Buehler, R., Ajello, M., et al. 2010, ApJ, 718, L166
Voges, W., Aschenbach, B., Boller, T., et al. 1999, VizieR Online Data Catalog, 9010, 0
Voges, W., Aschenbach, B., Boller, T., et al. 2000a, VizieR Online Data Catalog, 9029, 0
Voges, W., Aschenbach, B., Boller, T., et al. 2000b, VizieR Online Data Catalog, 9029, 0
Voges, W., Gruber, R., Haberl, F., et al. 1994, VizieR Online Data Catalog, 9010, 0
Watson, M. G., Schroder, A. C., Fyfe, D., et al. 2008, VizieR Online Data Catalog, 349, 30339