Probing the unidentified \textit{Fermi} blazar-like population using optical polarization and machine learning

I. Liodakis $^{1}$⋆ and D. Blinov $^{2,3,4}$

$^{1}$KIPAC, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA
$^{2}$Department of Physics and Institute for Theoretical and Computational Physics (ITCP), University of Crete, 71003 Heraklion, Greece
$^{3}$Foundation for Research and Technology – Hellas, IESL, Voutes, 7110 Heraklion, Greece
$^{4}$Astronomical Institute, St. Petersburg State University, Universitetsky pr. 28, Petrodvoretz, 198504 St. Petersburg, Russia

Accepted 2019 April 5. Received 2019 April 4; in original form 2018 December 28

ABSTRACT
The \textit{Fermi} $\gamma$-ray space telescope has revolutionized our view of the $\gamma$-ray sky and the high-energy processes in the Universe. While the number of known $\gamma$-ray emitters has increased by orders of magnitude since the launch of \textit{Fermi}, there is an ever increasing number of, now more than a thousand, detected point sources whose low-energy counterpart is to this day unknown. To address this problem, we combined optical polarization measurements from the RoboPol survey as well as other discriminants of blazars from publicly available all-sky surveys in machine learning (ML, random forest and logistic regression) frameworks that could be used to identify blazars in the \textit{Fermi} unidentified fields with an accuracy of $\sim$95 per cent. Of the potential observational biases considered, blazar variability seems to have the most significant effect reducing the predictive power of the frameworks to $\sim$80–85 per cent. We apply our ML framework to six unidentified \textit{Fermi} fields observed using the RoboPol polarimeter. We identified the same candidate source proposed by Mandarakas et al. for 3FGL J0221.2 + 2518.

Key words: methods: statistical – galaxies: active – galaxies: jets.

1 INTRODUCTION
The \textit{Fermi} $\gamma$-ray space telescope (\textit{Fermi}) has detected more than 3000 point sources since its launch in 2008 (Acero et al. 2015). The galactic population has been dominated by pulsars, while the vast majority of the extragalactic population (accounting for 58 per cent of the total number of associated sources) consists of blazars: active galactic nuclei (AGNs) with jets oriented close to our line of sight (Blandford, Meier & Readhead 2018). However, more than 1/3 of the total number of detected sources (as of the third \textit{Fermi} catalogue) are yet unassociated with a low-energy counterpart. This is of course due to \textit{Fermi}'s localization uncertainty which is of the order of a few arcminutes, much larger that e.g. a typical optical telescope.

The continuing observations of \textit{Fermi} seem to have two effects: (1) faint extragalactic $\gamma$-ray sources appear above the background (e.g. starburst galaxies, misaligned active galaxies); and (2) there is an ever increasing number of unidentified sources. The latter is often a consequence of the limited amount of follow up observations which is not nearly as much as what is needed to associate even a fraction of the unidentified population.

Typically, an AGN is associated with a detected point source in the \textit{Fermi} catalogue using either the spatial coincidence of the \textit{Fermi} point source with a source of a given catalogue (Bayesian association method, Abdo et al. 2010a) or the likelihood ratio association method which estimates an association probability based on the multiband (radio, optical, and X-ray) fluxes of sources within \textit{Fermi}'s positional uncertainty (Ackermann et al. 2015). Point sources not associated with a counterpart using these methods are labeled as unidentified. Several other methods have been attempted to identify AGN in the unidentified \textit{Fermi} fields (UFF). Popular methods include using the multiband properties of the sources (e.g. Acero et al. 2013); the jet morphology revealed through very long baseline array (VLBI) observations (Kovalev 2009); and infrared colours: blazars occupy a certain space in the infrared colour–colour diagram (e.g. Massaro et al. 2011). Most recently, Mandarakas et al. (2018) demonstrated that optical polarization can be a powerful probe for the $\gamma$-ray emitting AGN population. The method relies on the fact that $\gamma$-ray loud blazars have been found to be systematically more polarized than $\gamma$-ray quiet blazars (Angelakis et al. 2016). For sources above the galactic plane, where the interstellar dust induced polarization is expected to be less than 1–2 per cent, optical polarization could be used to confidently identify 544 $\pm$ 10 blazars and blazar-like sources still hiding in the UFFs (Mandarakas et al. 2018).

Although each method has its limitations, a combination of different probes could prove to be an effective tool in the fight against the increasing number of unassociated $\gamma$-ray emitting sources. In this work, we combine optical polarization observations with publicly available data from various all-sky surveys in machine learning (ML) frameworks and explore the potential of such tools.
In Section 2, we present the sample and observables used in this analysis, in Section 3, we built ML frameworks and evaluate their performance against observational biases, and in Section 4, we apply our framework to several unidentified fields of Fermi observed using the RoboPol polarimeter. We conclude in Section 5.

2 SAMPLE AND ANALYSIS

Our initial sample consists of γ-ray loud blazars observed in optical polarization by the RoboPol program. RoboPol observed a γ-ray flux-limited unbiased sample of blazars with high cadence for three years using a novel four-channel polarimeter (King et al. 2014). The sample was defined using strict statistical criteria (for details on the sample selection, see Pavlidou et al. 2014) making it suitable for statistical studies. The main goal of the RoboPol project was to explore the polarization properties of the jets and provide insight on the nature of the electric vector position angle (EVPA) rotations seen in blazars (Blinov et al. 2015, 2016a,b, 2018). Apart from the main sample (62 sources) that was monitored during those three years, roughly 200 in total γ-ray loud blazars were observed at least once as part of the so-called ‘June survey’ that was used for the final sample selection of RoboPol (Pavlidou et al. 2014). The sample observed as part of the June survey was also selected based on γ-ray flux (F(γ) > 100 MeV > 2 × 10^{-8} cm^{-2} s^{-1} in the second Fermi catalogue), optical magnitude (≤18), sky position (Galactic latitude |b| > 10), and elevation (E > 30 for at least 30 min in June). From all observed blazars, we selected those with at least three observations (93 sources). We complimented our sample with 11 additional γ-ray loud blazars observed by the Steward Observatory monitoring program (Smith et al. 2009). The total number of γ-ray loud blazars in our sample is thus 104 [hereafter γ-ray loud (GL) sample].

The RoboPol polarimeter allows for the simultaneous measurement of the Stokes q and u vectors without any moving parts. Each source on the sky is projected four times (four spots in a cross-like pattern) on the CCD camera each having its polarization orientation rotated by 67.5°. This allows for measurement of the polarization degree and EVPA of a source in a single exposure. For enhanced sensitivity and to block unwanted light, a mask is placed in the centre of the CCD, the location where the source of interest is typically positioned. A consequence of the RoboPol design is that sources outside the mask could have one (or more) of their spots overlap with a neighbouring source preventing us from confidently measuring either source’s polarization parameters (see fig. 6 in Panopoulou et al. 2015), especially in crowded fields. For every blazar observed with the RoboPol polarimeter, we use the automated reduction pipeline to extract the polarization degree for the sources in the 13 arcmin × 13 arcmin field of view of the RoboPol instrument excluding sources affected by such overlap (for details on the RoboPol pipeline and reduction process, see King et al. 2014; Panopoulou et al. 2015). This process yielded polarization measurements for 5837 sources. Additionally, we include 6379 sources from the Heiles (2000) catalogue of polarized stars residing above and below the galactic plane (|b| > 10° and |b| < −10°) resulting in a total of 12216 sources (hereafter γ-ray quiet (GQ) sample). Fig. 1 shows the median polarization degree of the GL and GQ samples. Clearly, the GL sources show systematically higher polarization due to the synchrotron nature of their emission as opposed to the, most likely, dust-induced polarization of the emission in GQ sources.

We use the Wide-field Infrared Survey Explorer (WISE) all-sky survey and the Two Micron All Sky Survey (2MASS, Skrutskie et al. 2006) to obtain the 3.4, 4.6, and 12 μm, and H-band magnitudes for all the sources in the GL and GQ samples. The WISE colours have been successfully used in the past to find blazars among the unidentified Fermi sources (e.g. Massaro et al. 2011). Fig. 2 shows the WISE colour–colour diagram along with the so-called blazar strip found in Massaro et al. (2011, 2012). While there is some minor overlap,

---

1 http://robopol.org/
2 http://james.as.arizona.edu/~psmith/Fermi/
3 Roughly, 27% of the sources are low synchrotron peaked, 20% are intermediate synchrotron peaked, 17% are high synchrotron peaked, and 36% are without synchrotron peak classification (Abdo et al. 2010b; Angelakis et al. 2016).
the GL sample occupies a very distinct area in the infrared colour–
colour diagram. We also estimate the $H$-band–4.6 $\mu$m colour. In
Majewski, Zasowski & Nidever (2011) stars were shown to have a
narrow distribution. The GQ sample (which is mostly comprised of
stars) shows a similar narrow distribution, whereas GL sources show
a wider distribution shifted to higher values (Fig. 3). In addition,
stars and nearby galactic sources could have measurable proper
motion while we do not expect any significant proper motion in distant
extragalactic sources (i.e. blazars). We use the Gaia DR2 catalogue
(Gaia Collaboration et al. 2016, 2018) to estimate the proper motion
(pm) and its uncertainty ($\sigma_{pm}$) for all the sources in the GL and GQ
samples. If the source has a $pm/\sigma_{pm} < 3$ we set $pm = 0$. Otherwise
we retain the estimated pm value. As expected, most of the sources
in the GL sample have pm consistent with zero. There are 10 sources
that show non-zero pm. This could be due to wrongful estimation of
the pm (e.g. fitting for pm did not converge) or underestimated $\sigma_{pm}$.
Whichever the case, we have verified that, for these 10 sources, using
the estimated pm values or setting them equal to zero does not have
any significant impact on the ML framework’s predictive power.

3 A MACHINE LEARNING APPROACH

The observables described above allow us to take advantage not
only of the intrinsic properties of the $\gamma$-ray emitting jets but also
the properties of the foreground sources and enable us to differentiate
them. Fig. 4 shows the relationship between all pairs of observables
for both GL and GQ samples. We first chose to combine them using
a random forest model (Breiman 2001). Random forests have been
increasingly used in astronomy, and successfully used in classifying
AGN sources of unknown type using their $\gamma$-ray properties (e.g.
Hassan et al. 2013; Saz Parkinson et al. 2016). They are a method
for classification by building a number of decorrelated decision
trees with randomly drawn input variables from the training set.
Each tree then casts a vote on the classification of the target source.
The majority vote is the final result. Random forests are built on
the principles of bootstrap aggregating (also known as bagging)
that reduce the variance of the model achieving better classification
results. We use the sklearn PYTHON implementation (Pedregosa et al.
2011) to set up a random forest framework initially consisting of 20
trees using 80 per cent of the GL and GQ samples for training and
20 per cent for testing. Fig. 5 shows the probability of classifying a
test source as GL for this particular choice for the number of decision
trees and training sample size. We varied the number of trees from
[20,200] with a 10 tree step as well as the percentage of sources in the
testing sample between 5–50 per cent to evaluate the stability of
the ML framework. In any combination of the above the accuracy
of the framework remains within $\leq 2$ per cent of the original set-up
($> 95$ per cent). To further assess the stability of the framework we
perform two tests. First, we create simulated training and testing
samples by randomly excising 30 per cent up to 70 per cent of the
initial samples and repeat the analysis. Even after $10^3$ trials, the
accuracy of the framework remained above 95 per cent. Second, we
artificially misclassified sources randomly in the training sample
starting from 1 per cent of the sample up to 100 per cent. We found
that if no more than 20 per cent of the sample is misclassified the
accuracy of the framework remains above 90 per cent.

An alternate approach which also serves as a validity check for
the random forest is a logistic regression model. Logistic regression
is a predictive regression method using a binary logistic model the
outcome of which can be either ‘success’ or ‘fail’. The logistic
model is defined as,

$$l = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n,$$

where $\beta_0$ is a constant value (intercept) and $[\beta_1, \beta_2, ..., \beta_n]$ are
the coefficients of the $[X_1, X_2, ..., X_n]$ input variables. The probability
for a given set of input parameters is given by,

$$p = \frac{1}{1 + e^{-l}}.$$

For $p > 0.5$ (or alternatively if $l > 0$) the outcome is ‘success’ or
in our case indicates a $\gamma$-ray emitting source. From the training
procedure we obtain $\beta_0 = -3.31$, and $[\beta_1, \beta_2,
\beta_3, \beta_4, \beta_5] = [1.45, 0.95, 0.74, -0.29, -2.67]$. For our
training $[X_1, X_2, X_3, X_4, X_5] = [\text{Pol. degree ( per cent)},
3.4 – 4.6 $\mu$m, 4.6 – 12 $\mu$m, H – 4.6 $\mu$m, pm]$. The logistic regression
yielded an accuracy of 95 per cent similar to the results of the
random forest.

Although in both cases, we achieve a high degree of accuracy, instru-
mental/observational systematics and/or the ever-present AGN
variability can affect the measurements of the parameters described
above hampering the accuracy of an already trained network. It is
then important to assess the stability of the predictive power of our
ML frameworks given the aforementioned effects.

3.1 Dependence on individual parameters

Individual measurements of the source parameters could be af-
fected due to a number of reasons. Although the dust-induced
polarization is expected to drop significantly above the galactic
plane, intrinsically polarized galactic sources (e.g. peculiar stars),
abnormally dusty regions in a given line of sight, or even erroneous
polarization measurements could result in non $\gamma$-ray sources having
comparable levels of polarization with the GL sample. In addition,
source confusion might affect crowded fields leading to erroneous
colour estimation. This effect might be particularly relevant for the
infrared observations given the larger point spread function of the
infrared telescopes compared to the optical. To assess the stability
of the ML framework to these effects we create simulated test samples
by drawing random values from the parameter distributions from the
GQ sample for all parameters but one. Since we are evaluating the
possibility of erroneous measurement or abnormally high values
we do not expect correlation between observations by different instruments. The parameter under investigation takes values from the minimum of the GQ sample to the maximum value of the GL sample with a step of roughly 1 per cent of the target variable. We then use the already trained network and, after $10^4$ simulations, we estimate the fraction of incorrectly identified sources. We only use sources from the GQ sample for this exercise since we want to assess the fraction of sources that could be misclassified as GL source, and because intrinsic variability in GL sources can induce much larger variations in the parameters. We explore the effects of intrinsic variability in the subsection below. Fig. 6 shows the results of this exercise for both ML frameworks. We find that both frameworks are most sensitive to the polarization degree and H–$4.6 \mu$m colour. In both cases, altering the H–$4.6 \mu$m colour results in a misclassification of only $\sim 2$ per cent at most. This is also true for the random forest and the polarization degree. However, the logistic regression shows a much higher dependence that could result in a misclassification of up to 80 per cent.

### 3.2 Blazar variability

Blazars are known to show violent and erratic variability throughout the electromagnetic spectrum and particularly in polarization (e.g. Angelakis et al. 2016; Kiehlmann et al. 2017). Given the large variations typically observed in the blazar light curves (e.g. Liodakis et al. 2018a,b), it is possible for the parameter values of a GL source
for each of the infrared magnitudes from a Gaussian distribution assuming a 20 per cent spread and recalculate the infrared colours. This results in a spread of 2–3σ typical for blazars. Once we have a simulated sample, we use the framework to predict whether the simulated sources are γ-ray emitters. We repeat this process 10⁴ and estimate the percentage of sources that where not correctly identified. The logistic regression seems to be less affected by variability with a median value of 15 ± 3.5 per cent misclassified sources whereas random forest shows a median of 22 ± 4 per cent. Although the relation between the polarization and flux variations in uncertain, blazars are known to flare simultaneously in multiple bands (e.g. Liidakis et al. 2017, 2018a). We repeat the above exercise twice. The first time we assume correlated variability between infrared colours, but the polarization degree varies randomly [correlated infrared (CI) case], while the second time all parameters are correlated [correlated infrared and polarization (CIP) case].

In both cases, the random forest framework yields similar results (20 ± 4 per cent for the CI and 19 ± 4 per cent for the CIP case). The same is true for the logistic regression (roughly 12 ± 3 per cent in both cases). Although increasing the spread of the parameters seems to increase the percentage of misclassified sources in all cases, for large values for the spread (> 100 per cent) there is a saturation at ̃35 per cent for the random forest and ̃45 per cent for the logistic regression. However, such extreme cases should be very rare given the fact that the most bright and variable blazars emitting γ-rays have most likely already been identified.

4 APPLICATION TO UNIDENTIFIED FERMI SOURCES

Given the high fraction of misclassifications by the logistic regression framework found in Section 3.1, for the application of our methodology, we opted to use the random forest framework. Mandarakas et al. (2018) used optical polarization to identify potential candidates for the γ-ray emitters in 4 UFFs. Several sources were observed within the 3σ positional uncertainty of each field (20 arcmin, Acero et al. 2015). Their analysis yielded a candidate source in the 3FGL J0221.2 + 2518 field at RA = 02h21m33.3s, Dec. = +25°12′47.5″. Follow-up spectroscopic observations confirmed its extragalactic origin while the comparison of the line ratios suggest a hybrid AGN-starburst galaxy. The source had not been included in published AGN catalogues and showed atypical, for a γ-ray emitting AGN, [3.4–4.6]μm, and [4.6–12]μm colours. The three other UFFs considered in Mandarakas et al. (2018) were 3FGL J0336.1+7500, 3FGL J0419.1+6636, and 3FGL J1848.6+3232. In our study, we include two additional fields also observed by RoboPol: 3FGL J0330.6+4037 and 3FGL J0855.4 + 7142. Each field was covered in three exposures. The first one was centred at the reported position of each UFF and the other two exposures were offset by 1 arcmin. Figs 7 and 8 show the positional uncertainty of Fermi (68 per cent and 95 per cent uncertainty, large circles) overlaid on a Sloan Digital Sky Survey (SDSS) image (Abolfathi et al. 2018). The small circles mark the sources observed by RoboPol.

We use the polarization estimates for all the sources observed in the UFFs and follow the same procedure as in Section 2 to estimate their infrared colours and proper motion. Fig. 9 shows the random forest probability of a source emitting γ-rays for all the sources observed by RoboPol. Only one source is classified as a GL source in 3FGL J0221.2 + 2518 with a 66 per cent probability marked with an arrow. The same source was proposed as the potential γ-ray emitter in Mandarakas et al. (2018). No
additional candidates were identified in any of the UFFs. However, that does not necessarily suggest that there are no blazar-like sources within those fields. Given the wide field limitations of the RoboPol instrument (Panopoulou et al. 2015, see also Section 2), it is likely that the source was either not observed or source overlap and/or other systematics prevented us from measuring its polarization degree (Figs 7 and 8). In addition, the limited exposure time used for the observations (~30–40 min per exposure) allows us to confidently estimate the polarization degree only for sources brighter than ~18 m in the R band. This prevents us from probing fainter sources that could be responsible for the γ-ray emission.

Based on the discussion in Section 3.2, it is also necessary to evaluate the probability of misclassifying blazar-like sources as non-γ-ray emitters. We use the ML probabilities of the simulated misclassified blazars in Section 3.2 to estimate the number of blazars with p-value between 0.1 (GQ sources show <0.1 p-values in the test sample, fig. 5) and 0.46 (the highest probability in the observed sample below the 0.5 threshold). We find that the majority of simulated misclassified blazars (~60 per cent) lay within that range suggesting that it is not unlikely for a misclassified blazar due to variability to exist within the observed fields. Additional observations are necessary to definitely exclude the possibility of a non-blazar-like source responsible for the γ-ray emission in those UFFs. Fig. 10 shows the infrared colours and polarization of all the sources observed in all six UFFs. The source in the top right corner marked with the vertical and horizontal lines, clearly standing out from the majority of the sources, is the candidate source found by the ML framework. In this particular UFF, for the next best candidate the random forest yields a p-value of 0.3. Only 19 per cent of the simulated misclassified blazars have probabilities between 0.1 < p-value <0.3.
5 CONCLUSIONS

We have demonstrated that optical polarization combined with other discriminants of jets and foreground sources in ML frameworks can be a powerful tool in the search for the missing blazar/blazar-like population in the still unidentified fields of Fermi. We introduced two new observables as predictors: the $H - 4.6\mu m$ colour and proper motion as measured by Gaia. Those allow us to effectively segregate nearby galactic sources from the extragalactic population where polarization can then easily distinguish synchrotron emission from e.g. dust-induced polarization or other mechanisms producing low degrees of optical polarization. Our approach is particularly useful for two main reasons. First, the majority of data used in this work are from publicly available, easily accessible, all-sky surveys. In addition, several upcoming polarization surveys (e.g. PASIPHAE, Tassis et al. 2018) will cover a large fraction of the sky providing measurements for the majority of sources in the UFFs. All of the above makes our proposed methodology among the least observationally expensive methods to find candidate extragalactic $\gamma$-ray emitters. Second, while individual observational probes have their limitations (e.g. the source in 3FGL J0221.2 + 2518 would not have been classified as a $\gamma$-ray emitter using the WISE colours), the combination of the observables discussed here limits their individual imperfections in identifying candidates. It should be noted that although 90 per cent of the GL sample is from the statistically complete unbiased RoboPol sample that describes the $\gamma$-ray emitting blazar population, the low number of sources used for the training of the framework could potentially affect its accuracy. Demographics could also be important, although there is no significant difference in the polarization degree (Angelakis et al. 2016) or proper motion between populations, and the blazars occupy distinct regions in the infrared colour space from other types of sources (e.g. Massaro et al. 2011). Additional measurements of $\gamma$-ray blazars (with an equally mixed composition of synchrotron peaked sources) would undoubtedly benefit our ML approach.

While confusion, erroneous measurements or other potential observational biases seem to have a small effect on the predictive power of the proposed ML frameworks ($<2$ per cent of sources could be misclassified for the random forest framework when varying a given parameter), blazar variability seems to affect it significantly (up to $\sim22$ per cent). Multiple measurements, additional predictors not affected (or less affected) by variability, or alternative ML approaches could help reduce this effect.

Applying our method to UFFs, our best candidate for 3FGL J0221.2 + 2518 is the same as the one proposed by Mandarakas et al. (2018). No additional candidate was found in five other fields without, however, excluding the possibility that the $\gamma$-ray emitter was not among the sources observed in polarization. Instruments built for wide field polarimetry such as WALOP$^5$ would remove this limitation by providing measurements for all the sources within the positional uncertainty of Fermi and could be used to effectively identify the $\gamma$-ray emitting sources in the UFFs. As noted in Mandarakas et al. (2018), there are two additional candidates for 3FGL J0221.2+2518: a radio galaxy NVSS J022126 + 251436 and a radio source detected by Schinzel et al. (2017). Although our approach disfavour NVSS J022126+251436 as the $\gamma$-ray emitter, we cannot exclude the possibility of that radio source emitting $\gamma$-rays, given that it was not detected during the polarization survey of 3FGL J0221.2 + 2518.

Our approach is in principle limited to blazars and blazar-like sources. However, application of the framework to the Fermi fields observed by Mandarakas et al. (2018) revealed that it can also be used to probe the non-traditional $\gamma$-ray AGN population. Polarization measurements of the increasing number of misaligned AGN detected by Fermi could allow us to build a more robust framework suitable for the entire extragalactic population of $\gamma$-ray emitters. Current and future large-scale surveys (e.g. eROSITA and LSST) could also help identify additional observables that could be easily incorporated into the proposed ML frameworks providing even more robust identifications for the entirety of the AGN $\gamma$-ray emitting population.

ACKNOWLEDGEMENTS

We thank G. Panopoulou and the referee T. Hassan for comments and suggestions that helped improve this work. DB acknowledges support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 771282. RoboPol is a collaboration involving the University of Crete, the Foundation for Research and Technology – Hellas, the California Institute of Technology, the Max-Planck Institute for Radioastronomy, the Nicolaus Copernicus University, and the Inter-University Centre for Astronomy and Astrophysics. Data from the Steward Observatory spectropolarimetric monitoring project were used. This program is supported by Fermi Guest Investigator grants NNX08AW56G, NNX09AU10G, NNX12AO93G, and NNX15AU81G. This publication makes use of data products from the WISE, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. This publication makes use of data products from the 2MASS, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. Funding for the SDSS IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório
I. Liodakis and D. Blinov

Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

REFERENCES

Abdo A. A. et al., 2010a, ApJS, 188, 405
Abdo A. A. et al., 2010b, ApJ, 716, 30
Abolfathi B. et al., 2018, ApJS, 235, 42
Acero F. et al., 2013, ApJ, 779, 133
Acero F. et al., 2015, ApJS, 218, 23
Ackermann M. et al., 2015, ApJ, 810, 14
Angelakis E. et al., 2016, MNRAS, 463, 3365
Blandford R., Meier D., Readhead A., 2018, preprint (arXiv)
Blinov D. et al., 2015, MNRAS, 453, 1669
Blinov D. et al., 2016a, MNRAS, 457, 2252
Blinov D. et al., 2016b, MNRAS, 462, 1775
Blinov D. et al., 2018, MNRAS, 474, 1296
Breiman L., 2001, Mach. Learn., 45, 5
Gaia Collaboration et al., 2016, A&A, 595, A2
Gaia Collaboration et al., 2018, A&A, 616, A1
Hassan T., Mirabal N., Contreras J. L., Oya I., 2013, MNRAS, 428, 220
Heiles C., 2000, AJ, 119, 923
Kiehlmann S., Blinov D., Pearson T. J., Liodakis I., 2017, MNRAS, 472, 3589
King O. G. et al., 2014, MNRAS, 432, 87
Kovalev Y. Y., 2009, ApJ, 707, L56
Liodakis I. et al., 2017, MNRAS, 466, 4625
Liodakis I., Romani R. W., Filippenko A. V., Kiehlmann S., Max-Moerbeck W., Readhead A. C. S., Zheng W., 2018a, MNRAS, 480, 5517
Liodakis I., Hovatta T., Huppenkothen D., Kiehlmann S., Max-Moerbeck W., Readhead A. C. S., 2018b, ApJ, 866, 137
Majewski S. R., Zasowski G., Nidever D. L., 2011, ApJ, 739, 25
Mandarakas N. et al., 2018, A&A, 623, A61
Massaro F., D’Abrusco R., Ajello M., Grindlay J. E., Smith H. A., 2011, ApJ, 740, L48
Massaro F., D’Abrusco R., Tosti G., Ajello M., Gasparini D., Grindlay J. E., Smith H. A., 2012, ApJ, 750, 138
Panopoulou G. et al., 2015, MNRAS, 452, 715
Pavlidou V. et al., 2014, MNRAS, 442, 1693
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Saz Parkinson P. M., Xu H., Yu P. L. H., Salvetti D., Marelli M., Falcone A. D., 2016, ApJ, 820, 8
Schinzel F. K., Petrov L., Taylor G. B., Edwards P. G., 2017, ApJ, 838, 139
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Smith P. S., Montiel E., Rightley S., Turner J., Schmidt G. D., Jannuzi B. T., 2009, preprint (arXiv:0912.3621)
Tassis K. et al., 2018, preprint (arXiv:1810.05652)

This paper has been typeset from a TEX/LATEX file prepared by the author.