Robustness of the Galactic Center excess morphology against masking

Yi-Ming Zhong[1,2,*] and Ilias Cholis[3,†]

[1]*Department of Physics, City University of Hong Kong, Kowloon, Hong Kong SAR, China*
[2]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*
[3]*Department of Physics, Oakland University, Rochester, Michigan 48309, USA*

The Galactic Center excess (GCE) remains an enduring mystery, with leading explanations being annihilating dark matter or an unresolved population of millisecond pulsars. Analyzing the morphology of the GCE provides critical clues to identify its exact origin. We investigate the robustness of the inferred GCE morphology against the effects of masking, an important step in the analysis where the gamma-ray emission from point sources and the galactic disk are excluded. Using different masks constructed from *Fermi* point source catalogs and a wavelet method, we find that the GCE morphology, particularly its ellipticity and cuspiness, is relatively independent of the choice of mask for energies above 2–3 GeV. The GCE morphology systematically favors an approximately spherical shape, as expected for dark matter annihilation. Compared to various stellar bulge profiles, a spherical dark matter annihilation profile better fits the data across different masks and galactic diffuse emission backgrounds, except for the stellar bulge profile from Coleman *et al.* Mon. Not. R. Astron. Soc. **495**, 3350 (2020)., which provides a similar fit to the data. Modeling the GCE with two components, one from dark matter annihilation and one tracing the Coleman bulge, we find this two-component model outperforms any single component or combinations of dark matter annihilation and other stellar bulge profiles. Uncertainty remains about the exact fraction contributed by each component across different background models and masks. However, when the Coleman bulge dominates, its corresponding spectrum lacks characteristics typically associated with millisecond pulsars, suggesting that it mostly models the emission from other sources instead of the GCE that is still present and spherically symmetric.

I. INTRODUCTION

The Galactic Center excess (GCE) has remained an unexplained component of gamma-ray emission for a decade and a half since it was first identified in the *Fermi*-LAT data by Ref. [1]. The GCE peaks in its energy output around a GeV. It appears to be a diffuse and extended component of gamma rays that cannot be explained by steady-state galactic diffuse emission, originating from the interaction with the interstellar medium of cosmic rays, typically produced by astrophysical sources such as pulsars or supernova remnants [2–14]. The exact origin of the GCE is still a topic of ongoing research debate, with the two most intriguing explanations being dark matter annihilation [7,9,15–20] and an unresolved population of millisecond pulsars (MSPs) [21–28]; while also a sequence of cosmic-ray burst events from the center of the galaxy provides a third alternative [29–31]. If the dark matter explanation is true, the GCE will not only serve as the first evidence that dark matter interacts with ordinary matter beyond its gravitational pull,

but it will also shed light on its production mechanism. If the MSPs explanation is true, it could reveal a new population of MSPs with luminosity function that is different from the known pulsars in globular clusters or the galactic disk [25,32]. Finally, if the GCE comes from a sequence of burst events, it would help us understand the conditions and recent history of the inner region of the Milky Way.

Various efforts have been pursued to identify the origin of the GCE, which include analyzing the details of its energy spectrum, its photon statistics, and its morphology [9,10,12,14,27,32–38] (see also [39]). A common step adopted in many GCE analyses is masking certain areas of the galactic center (GC). Masks are applied to exclude pixels of the sky where conventional sources of gamma-ray emission are expected to be so bright that even the uncertainty on their expected flux dominates the possible signal. Those include the galactic disk and the known point (or extended) sources that could interfere with the GCE signal. Various methods exist for creating such masks. One standard approach involves using the *Fermi*-LAT point source catalog to exclude regions around the cataloged sources, combined with band-shaped masks along $b = 0°$ to exclude the disk.

The choice of the mask and method to apply them to the data can potentially affect the results of the GCE analysis.

*yimzhong@cityu.edu.hk
†cholis@oakland.edu

For example, earlier works have demonstrated such masking effects on the non-Poissonian template fitting [23], where switching from masks created using the *Fermi*-LAT 3FGL catalog to the 4FGL catalog diminishes the statistical power of the method on the GCE photon statistics [27]. Refs. [14,32] investigated how the GCE characteristics are affected using alternative masks for the known *Fermi*-LAT point sources. They found its energy spectrum and morphology robust to the different updates of the *Fermi*-LAT catalog. Given the significance of the GCE morphology in its interpretation, we perform an extensive analysis of the robustness of the GCE morphology to alternative masking procedures in this work. We do that by combining with the vast publicly available library of galactic diffuse emission models produced in Ref. [14] to account for the astrophysical modeling uncertainties of the Milky Way conditions. Such an analysis is especially timely given the increasing number of point sources released in the latest *Fermi*-LAT catalogs [40–43].

We begin in Sec. II by introducing the *Fermi* data, the templates, the masks tested, and the statistical procedure used in our analysis. Section III presents our injection tests for the template fitting method in a controlled environment. Section IV shows the results of the GCE morphology under different masks. We find that the GCE morphology remains largely spherical and resembles that of a dark matter annihilation profile rather than that of a population of dim gamma-ray point sources tracing known stellar populations. An additional population of faint gamma-ray point sources tracing the stellar distribution known as the "Coleman bulge" [44] may exist. However, the spectrum associated with this component does not possess the characteristics expected from MSPs. We conclude in Sec. V.

II. SETUP FOR THE TEMPLATE FITTING

Here, we describe the *Fermi* data and the templates used for the diffuse emission and the GCE. We also describe the masks applied and the statistical procedure used in our template fitting analysis. Except for the mask construction, our analysis framework largely follows that of Refs. [14,38].

A. Data

We use *Fermi* Pass 8 data, version P8R3, recorded from August 4, 2008, to April 14, 2021, corresponding to weeks 9–670 of *Fermi*-LAT observations.[1] We use *Fermi* ScienceTools P8v27h5b5c8 for selection cuts and to calculate the relevant exposure-cube files and exposure maps,[2] which allow us to pass from fluxes to expected counts. The exposure is calculated for each pixel using the *Fermi* ScienceTools P8v27h5b5c8.

TABLE I. The energy bins and the energy-dependence of the fiducial radii θ_s (θ_l) used to mask known point sources with TS < 49 (TS ≥ 49).

	$E_{\min} - E_{\max}$ [GeV]	$\theta_s[°]$	$\theta_l[°]$
0	0.275–0.357	1.125	3.75
1	0.357–0.464	0.975	3.25
2	0.464–0.603	0.788	2.63
3	0.603–0.784	0.600	2.00
4	0.784–1.02	0.450	1.50
5	1.02–1.32	0.375	1.25
6	1.32–1.72	0.300	1.00
7	1.72–2.24	0.225	0.750
8	2.24–2.91	0.188	0.625
9	2.91–3.78	0.162	0.540
10	3.78–4.91	0.125	0.417
11	4.91–10.8	0.100	0.333
12	10.8–23.7	0.060	0.200
13	23.7–51.9	0.053	0.175

We keep only FRONT-converted CLEAN data. In addition, we set the following filters: zmax = 100°, DATA_QUAL == 1, LAT_CONFIG == 1, and ABS (ROCK_ANGLE) < 52. Our data maps are centered at the galactic center and cover a square window of 40° per side in galactic coordinates in Cartesian pixels of size 0.1° × 0.1°. Unlike a HEALPix pixelization, our pixels do not have equal area, but we account for this in our fits.

We bin the gamma-ray data in 14 energy bins spanning energies from 0.275 GeV to 51.9 GeV, given in the first column of Table I. The first eleven energy bins have a constant log width; because the gamma-ray flux drops at higher energy, the final three energy bins are wider, so each bin has a roughly similar statistical impact in our fits.

B. Galactic diffuse emission and GCE templates

We use the 80 astrophysical models that predict an equal number of sets of high-resolution energy-dependent templates developed in Ref. [14] for the galactic diffuse emission background.[3] Each set includes templates for π^0, inverse Compton scattering (ICS), and bremsstrahlung, calculated for each energy bin used in this analysis. In addition, we include an energy-independent template for the isotropic background and a template for the *Fermi*

---

[3]We use the term "background" to describe the combination of the galactic diffuse emission, the isotropic and mostly extragalactic emission, and the emission from the *Fermi* bubbles. If the GCE is the "signal," the term "background" is used in this context as the nonsignal emission. For our region of interest (ROI), depending on the latitude and energy, about 50–80% of the total emission and 60%–90% of the galactic diffuse emission is truly foreground emission, i.e., emission generated between our location and where the inner galaxy region begins (inner 3 kpc from the center of the Milky Way). See Ref. [10] for a more detailed analysis of this.

bubbles. In all cases, we projected the templates onto a $0.1° \times 0.1°$ Cartesian pixel grid near the galactic center that covers a square window of $40° \times 40°$. All templates are subsequently smoothed by the energy-dependent *Fermi*-LAT point spread function [45].

For the GCE, we consider contracted, squared, and integrated along the line-of-sight Navarro-Freak-White (NFW) profiles of the form,

$$\rho(r) = \frac{\rho_0}{(r/r_c)^\gamma (1 + r/r_c)^{3-\gamma}}, \qquad (1)$$

where the cuspiness parameter γ represents the inner slope of the density profile, $r_c = 20$ kpc is the scale radius and ρ_0 is the scale density that is fixed by $\rho(r = 8.5 \text{ kpc}) = 0.4 \text{ GeV/cm}^3$. We further deform the NFW profile by introducing an ellipticity (elongation) parameter ϵ, such that the opening angle from the GC, ψ, is related to the galactic coordinates (b, ℓ) by

$$\cos(\psi) = \cos(b) \cos(\ell/\epsilon). \qquad (2)$$

A spherically symmetric NFW template has $\epsilon = 1$ while an oblate (prolate) NFW template has $\epsilon > 1 (< 1)$. We construct the GCE template with a combination of γ and ϵ values. All the GCE templates are projected to the same Cartesian grid as the galactic diffuse emission and isotropic templates and smoothed by the same point spread functions.

C. Masks

We construct a variety of masks for the template fitting analysis.[4] Most of the masks are a combination of the point source mask and the disk mask.

Point source masks: We take the location information of the point sources from the 4FGL-DR1 [40], DR2 [41], and DR3 [42] catalogs. The point sources are classified into those with a test statistic (TS) < 49 and those with TS ≥ 49. We masked the two classes of point sources with disk-shaped masks that have energy-dependent radii θ_s and θ_l respectively (the same radii used in [14]), as listed in Table I. We call those point sources masks the "standard 4FGLDRX" with X = 1, 2, 3.

On top of the "standard 4FGLDRX" masks, we introduce two variations on the mask radii. For the "large 4FGLDRX" masks, we increase the fiducial radii by a factor of 1.5 for the point sources with TS < 49 while keeping radii for the sources with TS ≥ 49 unchanged. For the "small 4FGLDRX" masks, we decrease the fiducial radii by a factor of 0.5 for the point sources with TS < 49. The radii for the sources with TS ≥ 49 are again unchanged.

---

[4]The masks are available at https://github.com/ymzhong/gce_mask.

Disk masks: We set up three band-shaped masks for the galactic disk. For the standard disk mask ("L20"), we mask the region $|\ell| < 20°$ and $|b| < 2°$, i.e., the entire disk region of ROI. For the "L8" ("L5") disk mask, we mask the region $|\ell| < 8°$ ($|\ell| < 5°$) and $|b| < 2°$.

Combined regular masks: Combing the point source masks and the disk masks described above, we set up the following "regular" masks for our analysis: "standard 4FGLDRX + L20", "large 4FGLDRX + L20", "small 4FGLDRX + L20", "standard 4FGLDRX + L8", and "standard 4FGLDRX + L5". For "standard 4FGLDRX + L20", we create masks for X = 1, 2, 3. For other variations, the masks are only built for X = 2, 3. We show some of those combined masks in the first seven panels of Fig. 1.

Wavelet-based masks: The wavelet method [32,33] is a powerful way to identify potential point sources. Here, we utilize the wavelet peaks found in [32]. We select peaks with the statistic $\mathcal{S} > 2$, 3, and 4 (as defined in Refs. [32,33]), respectively, masking each peak with a disk-shaped mask that has energy-dependent radii identical to those used for the point sources with TS < 49. We then combine these wavelet peak masks with the "standard 4FGLDRX" point source masks and leave out the disk masks. The resulting masks are "standard 4FGLDRX + wavelet SY" with X = 2, 3 and Y = 2, 3, 4, where the Y values represent the thresholds of $\mathcal{S}$. We show the masks in the last three panels of Fig. 1.

For each mask, we compute the fraction of the masked pixels in each energy bin to the total number of pixels in the inner $40° \times 40°$ galactic center region (1.6×10^5 pixels). The results are shown in Fig. 2. Overall, the masked area decreases as the energy bin's energy increases. This reflects that the mask radii of the point sources decrease, following the *Fermi* point spread function as the energy increases. Looking into more detail, the top panel of Fig. 2 highlights the differences between masks created by different point source catalogs as well as by different radii criteria. As the number of point sources in the catalog increases, the masked pixel fraction significantly increases. The difference is significant for $E < 2.24$ GeV and gets diminished for higher energy bins as the number of masked pixels by the disk dominates over those by the point sources. Increasing or decreasing the radii criteria introduces a significant variation in the masked pixel fraction, which is the most significant for the lower energy bins. The middle panel of Fig. 2 highlights the difference in how we mask the galactic disk. Here, the difference between masking the entire galactic disk or parts of it is most significant at the higher energy bins ($E > 0.603$ GeV). For lower energy bins, the two sets of masks yield a similar masked pixel fraction, given that the point sources mask dominates the lower latitudes. Combining the information from the two panels, we can conclude that energy bins 0–3 are mostly affected by point sources masks, energy bins 8–13 are mostly affected by the disk mask, while energy
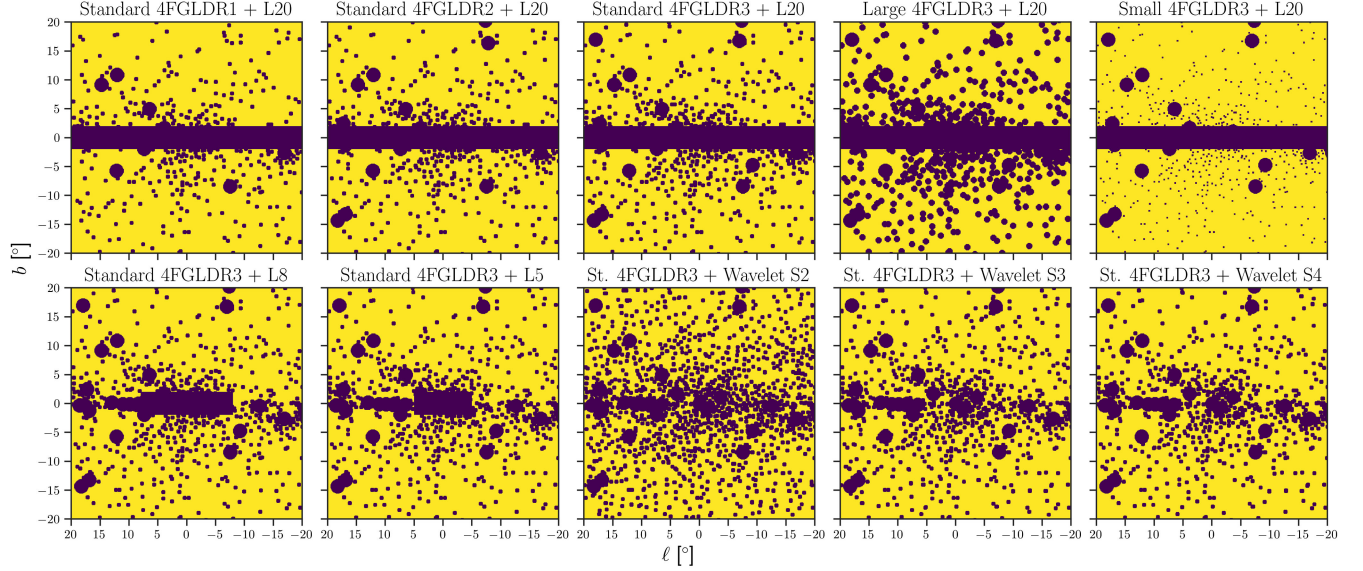
FIG. 1. The masks used in this work. Here, we show the masked regions (dark purple) for the energy bin 5 (1.02 GeV $< E <$ 1.32 GeV) only.

bins 4–7 are influenced by both types. The lower panel of Fig. 2 highlights the difference between different wavelet-based masks and those with a band mask for the disk. The wavelet-based masks, in general, mask a greater fraction of the GC region than the "standard 4FGLDRX + L20" masks at lower energy bins. However, as energy increases and the mask radii shrink, the wavelet-based masks are not blocked by the disk mask and have a sub 10% masked pixel fractions for $E > 3$ GeV. Also, at energies around 1 GeV, typically we only get 30% of the inner $40° \times 40°$ region masked.

FIG. 2. Masked pixel fraction for various masks used in the analysis. The vertical bands indicate the setup of 14 energy bins.

D. Statistical procedure

We construct the log-likelihood $\ln \mathcal{L}_j$ for a given energy bin j as,

$$
\begin{aligned}
-2 \ln \mathcal{L}_j = &\left\{ 2 \sum_p [\mathcal{M}_{j,p} \mathcal{C}_{j,p} + \ln [(\mathcal{M}_{j,p} \mathcal{D}_{j,p})!] \right. \\
&\left. - (\mathcal{M}_{j,p} \mathcal{D}_{j,p}) \ln(\mathcal{M}_{j,p} \mathcal{C}_{j,p})] \right\} \\
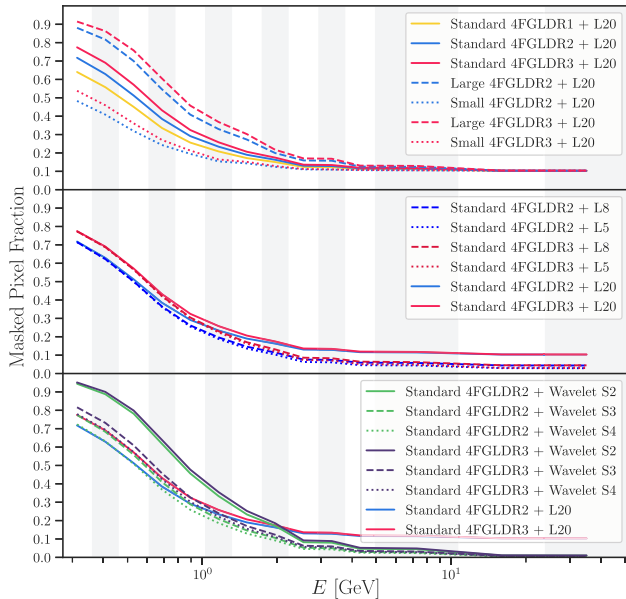&+ \chi^2_{\text{Bubbles},j} + \chi^2_{\text{Iso},j}.
\end{aligned}
\tag{3}
$$

The index p runs over the pixels of the ROI. The data $\mathcal{D}_{j,p}$ and the mask $\mathcal{M}_{j,p}$ are described in detail in Secs. II A and II C, respectively. The expected counts $\mathcal{C}_{j,p} = \mathcal{E}_{j,p} \sum_i c_j^i \Phi_{j,p}^i$ are obtained from summing over all the diffuse emission component templates $\Phi_{j,p}^i$ (i runs over different components) with independent normalizations c_j^i. We run fits without the GCE and with one or two templates to describe it. In all cases, our fits assume independent normalizations for each template. After linearly adding the templates, we multiply their sum by the exposure $\mathcal{E}_{j,p}$. The "external χ^2" functions $\chi^2_{\text{Bubbles},j}$ and $\chi^2_{\text{Iso},j}$ are constraints that act as penalties when the bubbles and isotropic normalizations deviate too much from their spectra measured at high latitudes, see Ref. [46,47].

We use EMCEE [48], a Markov chain Monte Carlo (MCMC) program, to sample the coefficient parameters c_j^i and to assess their posterior probability distribution. This process allows us to conveniently obtain the coefficient parameters that minimize the negative log-likelihood, Eq. (3), and identify the parameters' uncertainties. We use

TABLE II. The mock *Fermi* data and templates used for the injection tests. For the GCE templates, we use a spherical ($\epsilon = 1.0$) or oblate ($\epsilon = 1.4$) NFW profile with the cuspiness parameter of $\gamma = 1.2$. "GDE" stands for galactic diffuse emission model.

	Mock data	Fitting GCE template
(1)	$Fermi + 2 \times$ NFW ($\epsilon = 1.0$)	80 GDEs, NFW ($\epsilon = 1.0$)
(2)	$Fermi + 2 \times$ NFW ($\epsilon = 1.0$)	80 GDEs, NFW ($\epsilon = 1.4$)
(3)	$Fermi + 2 \times$ NFW ($\epsilon = 1.4$)	80 GDEs, NFW ($\epsilon = 1.0$)
(4)	$Fermi + 2 \times$ NFW ($\epsilon = 1.4$)	80 GDEs, NFW ($\epsilon = 1.4$)

ENSEMBLESAMPLER with 100 walkers and 1000 steps to assess convergence. We then discard the first 300 steps of the resulting chain from EMCEE and draw posterior probabilities for the parameters with CHAINCONSUMER [49]. The total

negative log-likelihood is a sum over the value of all the energy bins,

$$-2 \ln \mathcal{L} \equiv -2 \sum_j \ln \mathcal{L}_j. \quad (4)$$

III. INJECTION TEST

Injecting a GCE component into the *Fermi* data has been utilized to assess the sensitivity and robustness of a method in determining the characteristics of the GCE [20]. In this study, we conduct an injection test to evaluate how well the template fitting can identify the ellipticity parameter ϵ of the GCE. We investigate two GCE signals: a spherical NFW profile with parameters $\gamma = 1.2$, $\epsilon = 1.0$ and an
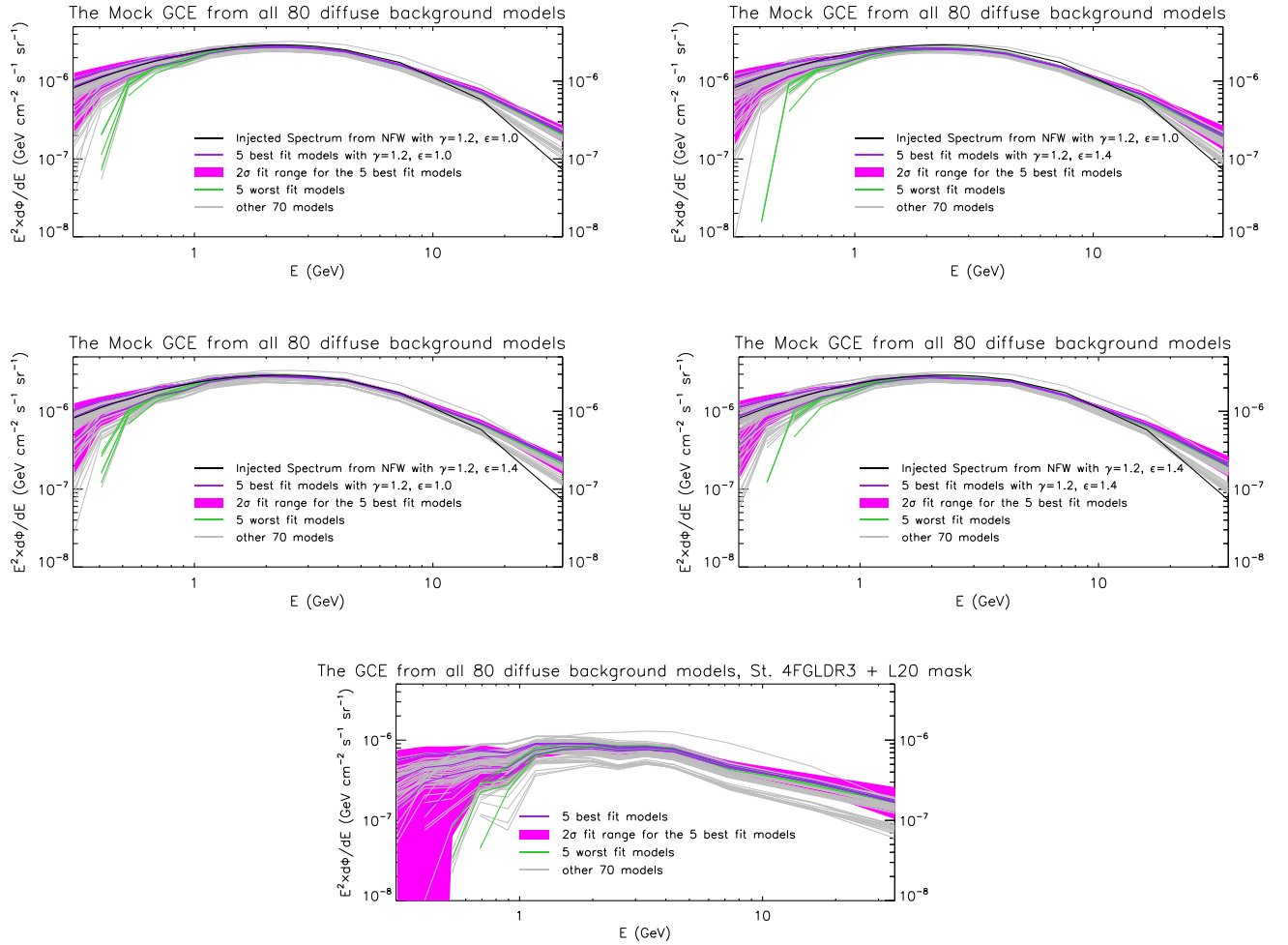

FIG. 3. Top two rows: The inferred energy spectra of the GCE from the mock *Fermi* data for the four injection tests, utilizing all 80 background galactic diffuse emission models of Ref [14] and the "standard 4FGLDR3 + L20" mask. Further details are provided in the text. Bottom panel: the inferred GCE spectrum from the actual *Fermi* data for comparison. In the fit, the GCE flux is permitted to be negative, although this occurs in only a small subset of models and exclusively at energies below 0.7 GeV (indicated by the termination of lines). For the five background models out of the 80 that yield the best overall fit to the gamma-ray data, the best-fit GCE normalization and its corresponding 2σ uncertainty range are represented by purple lines and magenta bands, respectively. The green lines give the GCE normalization from the five galactic diffuse emission models that provide the worst fit to the gamma-ray data out of the 80 considered.
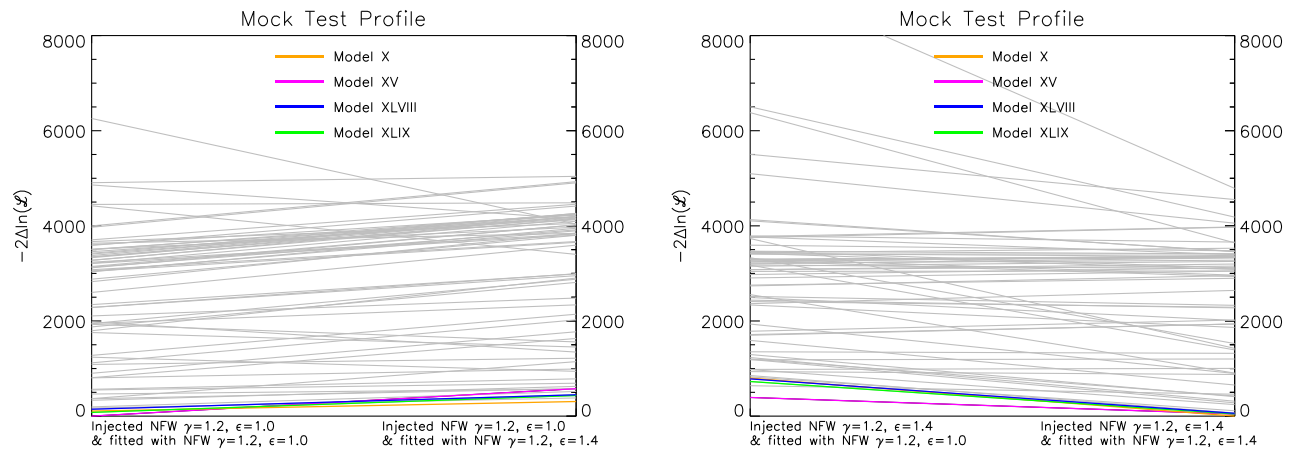
FIG. 4.   Comparison of the difference in the total negative log-likelihood, $-2\Delta \ln \mathcal{L}$, for injection test scenario (1) vs scenario (2) (left), scenario (3) vs scenario (4) (right). Colored lines represent some of the galactic diffuse emission models that give the best fit to the *Fermi* data.

oblate NFW profile with $\gamma = 1.2$, $\epsilon = 1.4$. The GCE signals are injected into the *Fermi* data with a flux twice the fiducial value. Note that the template fit of the actual *Fermi* data indicates that the observed flux of the GCE is around the fiducial values for both signals. Utilizing these mock *Fermi* datasets, we conduct four separate template-fitting analyses as outlined in Table II. In these tests, scenarios (1) and (4) match the injected data with the correct GCE template, whereas scenarios (2) and (3) involve a mismatch between the injected data and the template.

The resulting energy spectra of the GCE are shown in the top two rows of Fig. 3, where we apply the "standard 4FGLDRX + L20" mask. The magenta bands indicate the 2σ fit ranges for the five best-fit background galactic diffuse emission models. For comparison, the lower panel of Fig. 3 shows the GCE spectrum obtained from fitting the actual *Fermi* data, applying the same mask. Upon initial examination, all the injection tests, regardless of whether the GCE template matches or mismatches, show fitted spectra that are consistent with the injected GCE spectra (shown as black lines) up to 10 GeV [except a small discrepancy in scenario (2)]. Notably, the fluxes in these tests are approximately three times higher than those obtained from the actual *Fermi* data. That should be the case since the *Fermi* data already contains the GCE component before injecting the additional GCE mock component into it.

We further look into the specifics by comparing the total negative log-likelihoods in scenario (1) vs scenario (2) and scenario (3) vs scenario (4). The results are shown in Fig. 4 in terms of the difference of total negative log-likelihoods, $-2\Delta \ln \mathcal{L}$, relative to the best-fit models that include both the galactic diffuse emission backgrounds and the GCE. Our analysis shows that, in general, the cases where the GCE model matches the injected data [scenarios (1) and (4)] result in a better fit compared to their mismatched counterparts [scenarios (2) and (3)] across the majority of

the 80 galactic diffuse emission background models fitted. The difference in $-2\Delta \ln \mathcal{L}$ ranges from approximately 50 to 100 for the five best-fit background models when comparing matched with mismatched scenarios. Only a few background models that yield poorer fits show a preference for the mismatched scenarios over the matched ones. Our results suggest that the template fitting demonstrates robustness against misinterpreting the ellipticity of the GCE and quantifies the relevant statistical penalty.

IV. THE MORPHOLOGY OF THE GCE

In the subsequent sections, we provide a comprehensive list of tests focusing on the morphology of the GCE. Our investigation covers the properties of the GCE morphology, i.e., its cuspiness and ellipticity, across different energies and under different masks as introduced in Sec. II C (see Fig. 1). Additionally, we evaluate and compare various alternative models for the GCE morphology proposed in the literature. These models encompass scenarios such as a dark matter annihilation profile, a series of stellar bulge profiles, and a combination of the dark matter annihilation profile and a stellar bulge. Each model is examined under different masks to determine their respective goodness of fit.

A. The dependence of the GCE morphology with energy

One key question in determining the origin of the GCE is understanding the dependence of its morphology on energy. We expect an energy-independent morphology if the GCE originates purely from dark matter annihilation. This scenario assumes that gamma rays from the underlying annihilation channel are prompt emission photons rather than (ICS) photons up-scatted from high-energy electrons and positrons that interact with the interstellar medium as they propagate away from their point of origin.

Dark matter particles give a dominant prompt emission component when they annihilate mainly into quarks and gluons. Instead, if the GCE is purely of stellar origin, it would consistently show an elongation along the Galactic disk ($\epsilon > 1$), regardless of energy. Such a morphology is consistent with predictions from stellar bulge profiles, where the gamma-ray emission is predominantly from the outer gap regions of MSPs and thus directly tracks the location of the MSPs. Finally, the *Fermi* bubbles might significantly affect the observed GCE emissions, or even be related to it, being an earlier event of cosmic-ray burst activity from the GC [29–31]. Given the hard spectrum of the *Fermi* bubbles, their contribution to the GCE is expected to increase with energy. In this scenario, the GCE will exhibit a more pronounced elongation perpendicular to the galactic disk at higher energies, i.e., give ellipticity parameter values of $\epsilon < 1$.

In Fig. 5, we show how the GCE morphological properties change with energy. The left panel, assuming the GCE is spherical, shows the evolution of the cuspiness parameter γ from Eq. (1) with energy. This parameter describes the level of contraction in the NFW profile. We remind the reader that the GCE morphology is proportional to the line-of-sight integral of $\rho(r)^2$. We use mask "standard 4FGLDR3 + L20" and include results for all 80 background galactic diffuse emission models from [14] and highlight in color the results coming from background galactic diffuse emission models that provide the best fit to the *Fermi* data. At low energies, the systematic uncertainties of the background diffuse emission models prevent a definitive conclusion about the value of γ. We remind the

reader that various alternative background models can give similar quality fits. However, for gamma-ray energies above 2 GeV, there is a robust preference for the GCE to exhibit a cuspiness of $\gamma \geq 1$, akin to a regular NFW profile or a slightly adiabatically contracted version. Only a minority of background models at high energies favor a cuspiness of $\gamma < 1$, but these models generally have poorer overall fit quality for the ROI. In Appendix A, we provide the results from the same analysis for alternative ROIs (masks). For the regular masks, our result on the cuspiness parameter being $\gamma \simeq 1.2$ remains valid. It should be noted that, in addition to the systematic uncertainties of the galactic diffuse emission backgrounds, accurately tracking the GCE morphology at lower energies is challenging due to the significant fraction of the ROI being removed by point source masks, as detailed in Sec. II C and Fig. 2.

In the right panel of Fig. 5, again using "standard 4FGLDR3 + L20" mask, we show how the ellipticity parameter ϵ of Eq. (2) changes with energy. This parameter indicates the sphericity of the GCE. Again, we show results from all 80 background diffuse emission models of Ref. [14] and highlight in color those from the best-fit background models. Similar to the cuspiness analysis, at low energies, the combination of large astrophysical uncertainties of backgrounds and the small fraction of the unmasked ROI prevents us from conclusively determining a preference for ϵ. Nonetheless, at gamma-ray energies $\gtrsim 3$ GeV, the background models that provide a good fit to the data consistently suggest that the GCE is *approximately* spherical. In Appendix A, we show results on the ellipticity parameter for alternative masks. An ellipticity $\epsilon \simeq 1.0$–1.4
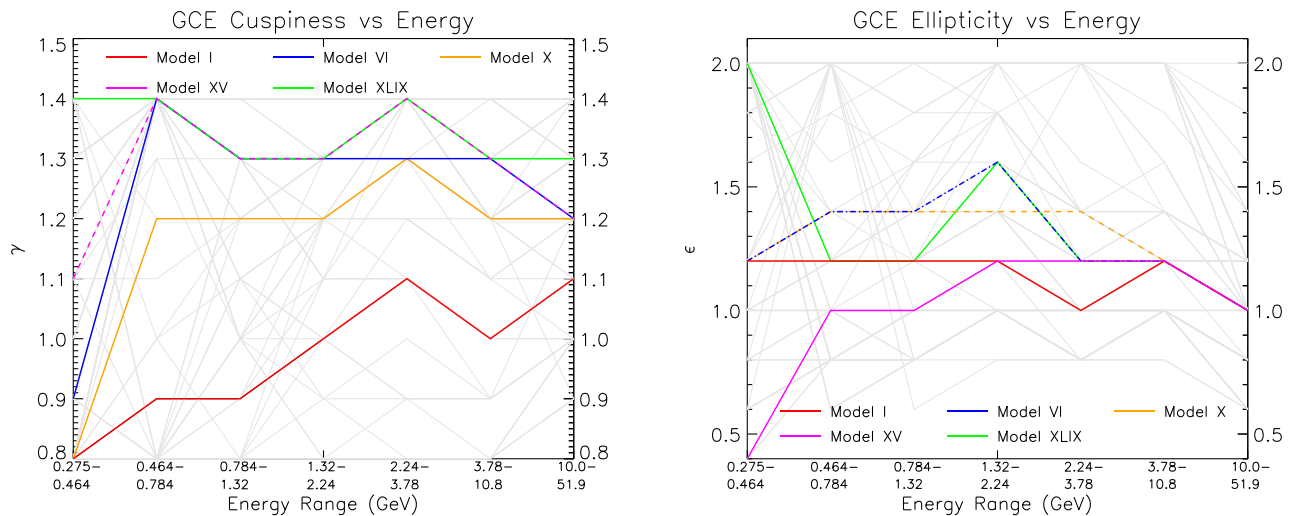

FIG. 5.   The evolution of the fit GCE morphology with energy using the "standard 4FGLDR3 + L20" mask. Left: how the inner slope (cuspiness) parameter γ changes with energy. We show the results for all 80 galactic diffuse background models in gray lines, where we show what the best-fit γ value was at a given energy range for each model. With the colored lines, we show the same results for some of the best overall fit models for the entire energy range. Right: Similarly, we show how the ellipticity parameter ϵ changes with energy. At low energies, due to the large masks and astrophysical background uncertainties, the GCE morphology, both in terms of its cuspiness γ and ellipticity ϵ, cannot be well constrained. However, with increasing energy as the point sources mask shrinks, the best-fit models converge to $1.1 \leq \gamma \leq 1.3$ and $\epsilon \simeq 1.0$. Some background models that provide a poor fit to the data still allow for a wider range of γ and ϵ.

for energies $\gtrsim 3$ GeV is still the case for the regular masks. The mild derivation from a perfect spherical shape ($\epsilon = 1.0$) could be due to the baryonic contraction of the dark matter halo [50].

B. The cuspiness and ellipticity of the GCE for alternative masks

In this section, we test the general morphological properties of the GCE in terms of its cuspiness and ellipticity under various mask choices.

Figure 6 examines the cuspiness γ of the GCE. We show the quality of fit by comparing differences in total negative log-likelihood, $-2\Delta \ln \mathcal{L}$. This statistic reflects how well the combination of an astrophysical background and GCE model agrees with the *Fermi* data for each of the six masks presented. To encompass the astrophysical uncertainties of backgrounds, we analyze all 80 astrophysical galactic diffuse emission models provided in Ref. [14]. In our plots of $-2\Delta \ln \mathcal{L}$, we set baseline (zero point) as the combination of background and GCE model that yields the best overall fit once summing over all energy bins. Lines closer to zero indicate a higher quality fit. Since we plot the $-2\Delta \ln \mathcal{L}$ values up to 8000, in some of the panels (especially the lower ones of Fig. 6), only a subset of lines is visible. The absent lines are models with poorer fits. For each panel, we color the combination of background and

GCE model that provides the best fit, along with four more models that provide comparatively similar quality fits.

Direct comparison of log-likelihoods across different masks is not feasible, as the number of pixels in the remaining ROIs where fitting occurs varies. Also, the exact ranking of which combination of galactic diffuse emission background models and GCE provides the best-fit changes slightly with different ROIs (masks). This variation is expected, mainly because regions dense in interstellar medium gas may be either revealed or masked by altering the masks, and these regions are typically associated with high gamma-ray fluxes. Nonetheless, our objective is to identify the properties of the GCE morphology that are robust regardless of these varying assumptions. We achieve this by searching for recurring patterns across the combination of different choices in the diffuse emission backgrounds and masks.

In the case of Fig. 6, a clear pattern emerges: irrespective of the background model and the mask used, the GCE consistently shows a preference for $1.0 \leq \gamma \leq 1.4$. This range corresponds to the morphology of a dark matter annihilation signal, following a spatial distribution akin to a regular NFW profile ($\gamma = 1$) or a more contracted profile that scales as $\sim 1/r^\gamma$ at the inner galaxy.

These findings are not affected by using any of the remaining four masks (ROIs) of Fig. 1: "standard 4FGLDR1 + L20," "standard 4FGLDR2 + L20," "standard
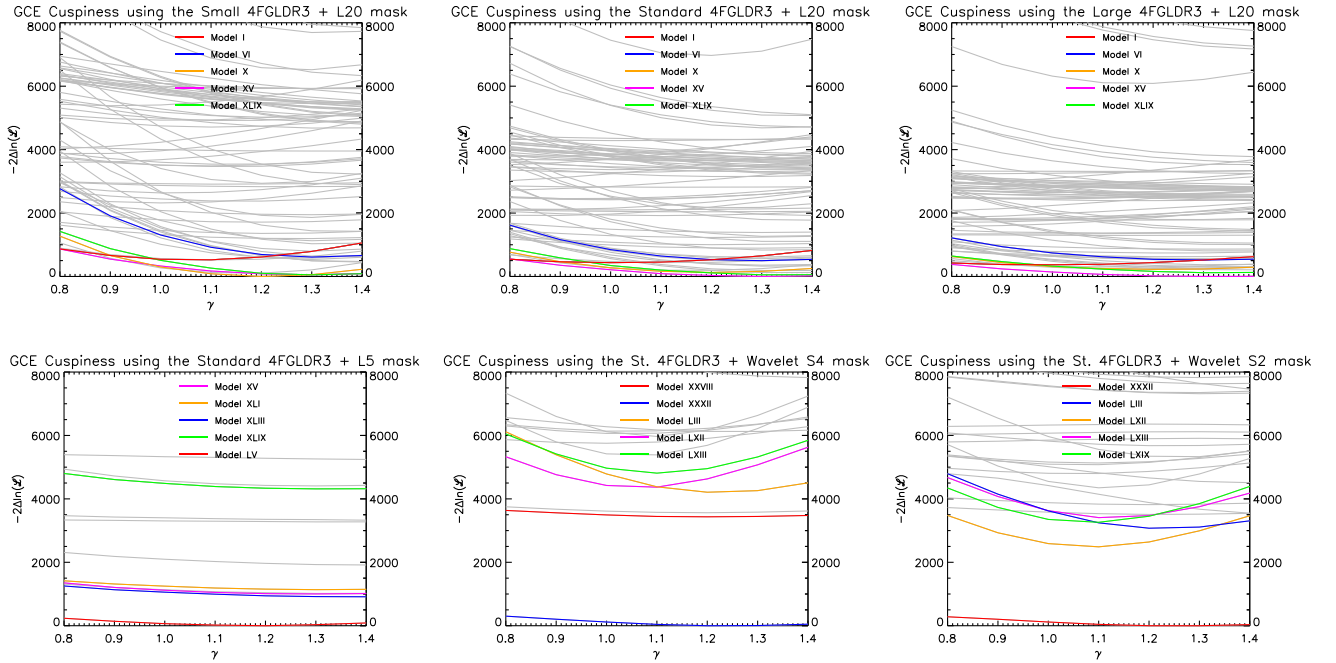

FIG. 6. The GCE cuspiness assuming the emission comes from annihilating dark matter. Log-likelihoods are evaluated by summing all energy bins. The dark matter profile is described by the γ parameter of Eq. (1). We show results for alternative mask choices. From the top left panel to the bottom right, we use masks, "small 4FGLDR3 + L20," "standard 4FGLDR3 + L20," "large 4FGLDR3 + L20," "standard 4FGLDR3 + L5," "standard 4FGLDR3 + wavelet S4" and "standard 4FGLDR3 + wavelet S2." The alternative masks (ROIs) may affect the ordering between the diffuse emission background models that provide the best fit and the difference in $-2\Delta \ln \mathcal{L}$, but the common pattern is a systematic preference for $1.0 \leq \gamma \leq 1.4$.
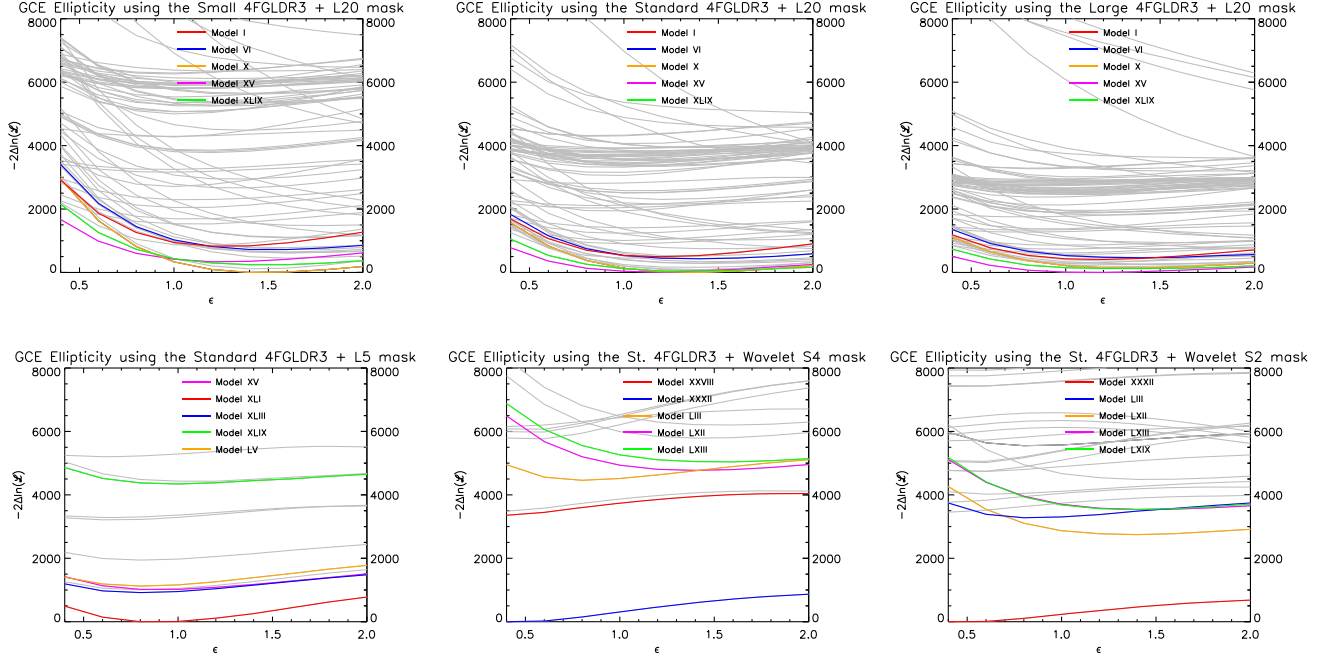
FIG. 7. As in Fig. 6, we show the GCE ellipticity for alternative choices of masks. Log-likelihoods are again evaluated by summing all energy bins. From the masks (ROIs) shown in the top panels, the backgrounds giving the best fit prefer $0.9 \leq \epsilon \leq 1.5$. Using instead as a mask the "standard 4FGLDR3 + L5", we get $0.7 \leq \epsilon \leq 1.2$. The two wavelet-based masks of "standard 4FGLDR3 + wavelet S4" and "standard 4FGLDR3 + wavelet S2" give for a small number of models, but also the ones providing the best fit, a preference for $0.4 \leq \epsilon \leq 0.8$.

4FGLDR3 + L8," "standard 4FGLDR3 + wavelet S3." The "standard 4FGLDR1 + L20" and "standard 4FGLDR2 + L20" masks were tested in Ref. [14], yielding the same conclusions as in this work. The "standard 4FGLDR3 + L8" and "standard 4FGLDR3 + wavelet S3" masks represent intermediate masking assumptions compared to those plotted here.

Figure 7 focuses on examining the ellipticity ϵ of the GCE. Again, we show the quality of fit in terms of $-2\Delta \ln \mathcal{L}$ for the combined astrophysical background and GCE models against the *Fermi* data, applying the same six masks as in Fig. 6. Lines closer to zero in $-2\Delta \ln \mathcal{L}$ represent better quality fits. In each case, we color the combinations of background and GCE models that provide the best fit, along with four additional background models yielding similar agreement to the observations.

The ellipticity of GCE appears to be more sensitive to the combination of assumptions used, compared to its cuspiness. Under the masks shown in the top panels ("small 4FGLDR3 + L20," "standard 4FGLDR3 + L20," and "large 4FGLDR3 + L20"), the diffuse emission background models that give the best fit prefer $0.9 \leq \epsilon \leq 1.5$. The "standard 4FGLDR3 + L5" mask yields a similar range of $0.7 \leq \epsilon \leq 1.2$. However, wavelet-based masks, namely "standard 4FGLDR3 + wavelet S4" and "standard 4FGLDR3 + wavelet S2", give a range of

$0.4 \leq \epsilon \leq 0.8$ for a few models that provide a better fit. Such values of ϵ suggest that the GCE is elongated perpendicular to the galactic disk, which would indicate the GCE comes from a burst of cosmic rays similar or the same as the one given in the *Fermi* bubbles. Such an ellipticity may also due to a triaxial dark matter profile [51–54]. However, this ellipticity is unlikely to result from a population of gamma-ray sources tracing the dense stellar regions of the Milky Way. Additionally, it could also be due to a nontrivial selection effect in which pixels are identified by the wavelet method as potential point sources.

C. Comparing dark matter annihilation and stellar bulge profiles

In this section, we compare the hypothesis of dark matter annihilation being responsible for the GCE with the alternative theory of unresolved gamma-ray point sources, such as MSPs, being the origin. The latter hypothesis suggests that the GCE, to some extent, follows one of the known stellar populations, such as the galactic bulge or the nuclear stellar cluster (nuclear bulge) at the center of the Milky Way. For the dark matter annihilation hypothesis, we adopt the NFW profile with $\gamma = 1.2$ and $\epsilon = 1.0$ (referred to as "NFW $\gamma = 1.2$"), as this choice typically provides the best fit to the gamma-ray data. We compare this with a sequence of alternative morphological models proposed in the literature for the GCE.
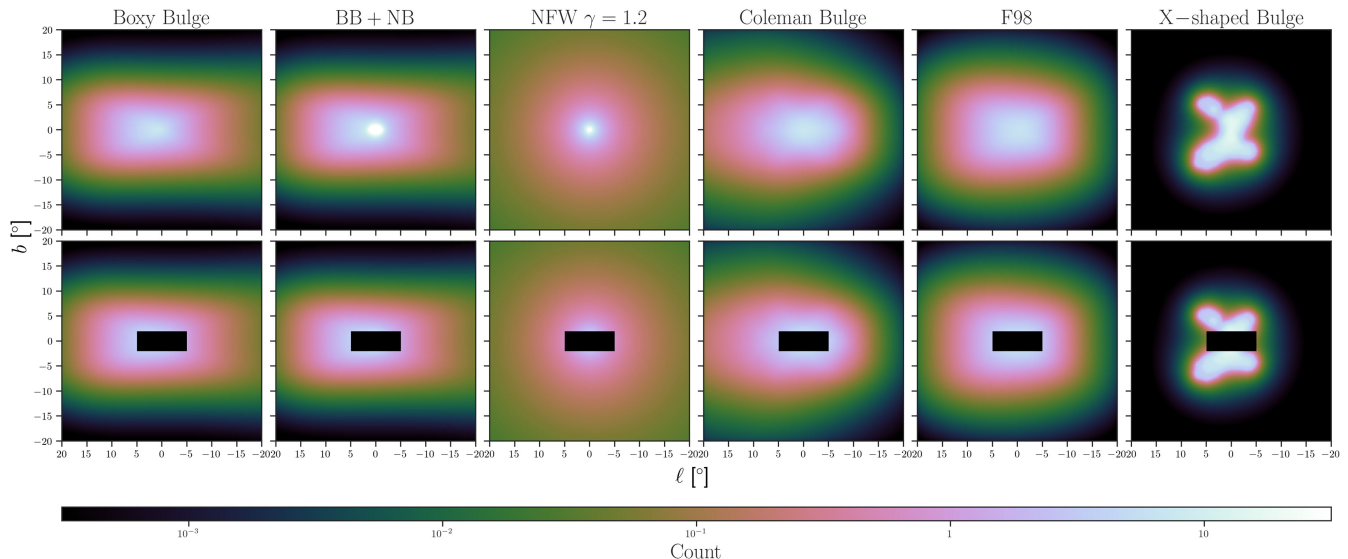
FIG. 8.   The boxy bulge (1st col.), BB + NB (2nd col.), Coleman bulge (4th col.), F98 (5th col.), and X-shaped bulge (6th col.) profiles used in this work, together with the NFW $\gamma = 1.2$ profile (3rd col.), for energy bin 5 (1.02 GeV $< E <$ 1.32 GeV). The top row shows the profiles in the full 40° × 40° ROI. The bottom row shows the same profiles with the "L5" disk mask, i.e., masking the region with $|\ell| < 5°$ and $|b| < 2°$. All the profiles are smeared by the *Fermi* PSF. The color represents the photon counts of the fiducial profiles used in the template fitting.

We compare the dark matter annihilation template and the following stellar bulge profiles:

(i) The "boxy bulge" profile, based on the stellar bulge measured at the infrared [55].

(ii) A combination of the boxy bulge and the nuclear bulge "BB + NB". The nuclear bulge is measured at radio waves [55]. This combination has been proposed as the correct morphology for the GCE in Ref. [26].

(iii) The Coleman bulge from Ref. [44], which has been recently suggested as an alternative for the GCE morphology [56].

(iv) The Freudenreich profile for the galactic bar, "F98." It is measured at the infrared [57]. This profile has been tested for the GCE morphology in Ref. [34] and recently suggested as a possible alternative for the GCE morphology [56].

(v) The "X-shaped bulge," observed in lower energy gamma rays and suggested as having the correct morphology for the GCE in Refs. [34,35].

We showcase the morphology of the stellar bulge profiles, together with the NFW $\gamma = 1.2$ profile, in Fig. 8. The top row shows the photon count maps, smeared by the *Fermi* PSF, in the 1.02–1.32 GeV energy bin (energy bin 5) of the fiducial templates. The bottom row shows these same spatial maps after applying the "L5" disk mask, which covers the brightest regions across all the profiles and is the largest common masking area used in the regular masks. With the L5 disk mask applied, the Coleman bulge profile bears the greatest similarity to the spherical NFW $\gamma = 1.2$ among all the stellar bugle profiles.

The results from template fitting are shown in Fig. 9. We find that the dark matter annihilation template, NFW $\gamma = 1.2$, generally outperforms the boxy bulge, BB + NB and the X-shaped bulge morphologies, irrespective of the mask or the astrophysical background model for the galactic diffuse emissions.[5] These findings strengthen previous results from Ref. [14], which reached similar conclusions regarding the GCE morphology using the standard 4FGLDR1 + L20 and standard 4FGLDR2 + L20 masks.

Interestingly, the F98 model's performance compared to the NFW $\gamma = 1.2$ varies depending on the specific combination of mask and background models. The difference in the fit quality between the two models for the GCE morphology is typically small. However, the NFW $\gamma = 1.2$ morphology still slightly outperforms the F98 profile, particularly with regular masks such as standard 4FGLDR3 + L20, small 4FGLDR3 + L20, large 4FGLDR3 + L20, and standard 4FGLDR3 + L5, for the best-fit background choices. The preference between NFW $\gamma = 1.2$ and F98 varies more with wavelet-based masks, but NFW $\gamma = 1.2$ still outperforms the best-fit background choice.

---

[5]For some background models the boxy bulge or BB + NB profiles may seem to provide a better than NFW $\gamma = 1.2$. However, these cases involve background models that yield a significantly poorer quality of fit. It highlights the importance of background model selection, as a suboptimal choice can potentially give misleading conclusions about the morphological properties of the GCE (see also Ref. [38]).

When comparing the NFW $\gamma = 1.2$ to the Coleman bulge profile, we find that the exact choice of mask and galactic diffuse emission background affects the preference between the two GCE templates. However, the difference in $-2\Delta \ln \mathcal{L}$ is typically small. As shown in the fourth column of Fig. 8, the Coleman bulge profile has a morphology that is asymmetric peanut shape, elongated along the galactic disk. However, when comparing to other stellar bulge profiles, we note that *the Coleman bulge profile is morphologically the closest to a spherical NFW profile among the bulge profiles*, particularly when masking the central L5 region. Such similarity makes the fit result less surprising. We also note that which of those (if any) bulge profiles more faithfully represents the population of stellar progenitors to the possible MSPs laying below the point source threshold remains an open but important question.

Furthermore, we explore whether the GCE could be a combination of dark matter annihilation and a stellar bulge profile, with free relative normalization for each of the 14 energy bins we study. This can be justified as even if the GCE is mostly a dark matter signal, we expect contribution from unresolved MSPs and other gamma-ray sources at some level. We test combinations of the NFW $\gamma = 1.2$ profile with (1) the BB + NB profile, (2) the F98 profile, and (3) the Coleman bulge profile. For the first two cases,

the NFW $\gamma = 1.2$ absorbs almost all the GCE-associated emission above 0.7 GeV effectively in all background model combinations and masks. Furthermore, the difference in log-likelihood between using the combined profile and the pure NFW $\gamma = 1.2$ template is small. That is in agreement with earlier results when we used the "standard 4FGLDR2 + L20" mask in Refs. [14,38].

However, the combination of NFW $\gamma = 1.2$ with the Coleman bulge profile (NFW $\gamma = 1.2$ & Coleman bulge) presents a less definitive picture. Practically in all scenarios, the combined profile NFW $\gamma = 1.2$ & Coleman bulge performs better or as well as either the NFW $\gamma = 1.2$ or the Coleman bulge profiles alone. The ranking again varies depending on the mask and the background model. This leads us to explore further the amount of the GCE emission that is absorbed by each of the two components (NFW $\gamma = 1.2$ or Coleman bulge) as a function of energy.

D. The GCE spectrum from dark matter annihilation and a population of point sources

In this section, we focus on the case that the GCE is attributed to two overlapping components, each with its distinct morphology. One component is modeled by the spherically symmetric dark matter annihilation template,
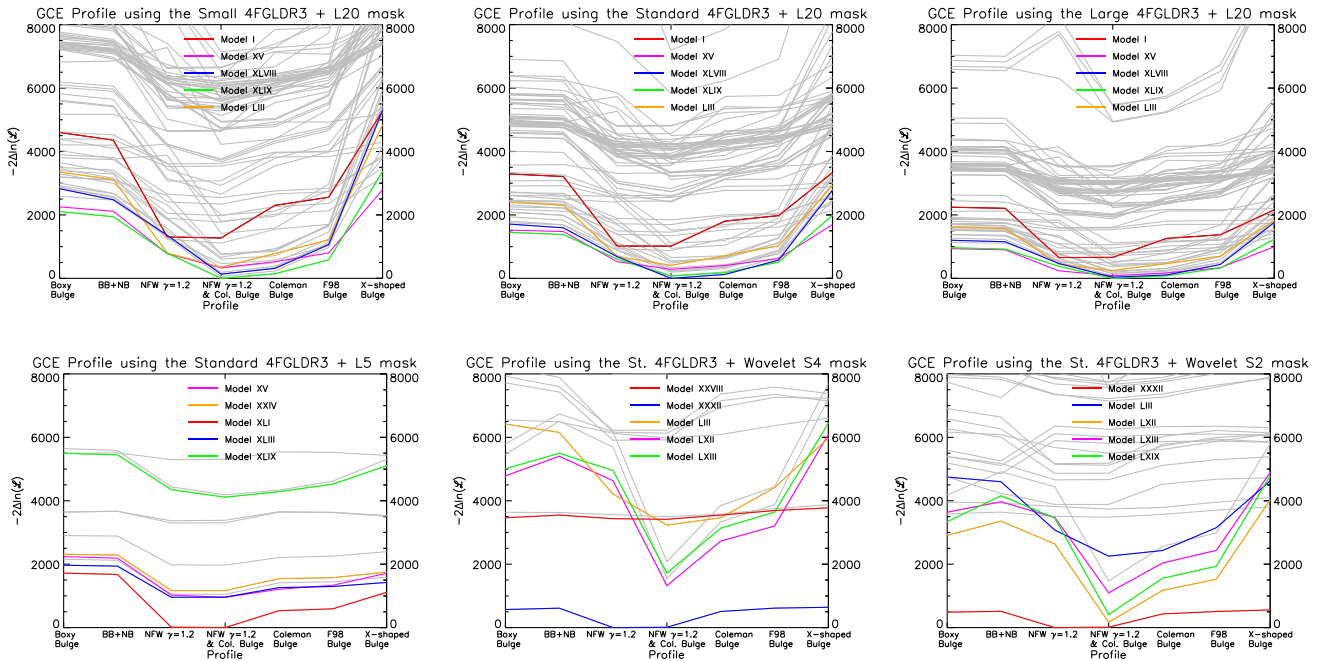

FIG. 9. Similarly to Figs. 6 and 7, for alternative choices of masks we show how the assumption that the GCE comes from dark matter annihilation (NFW $\gamma = 1.2$) compares to various choices of bulges for the GCE (see text for details). Log-likelihoods are again evaluated by summing all energy bins. We also include the option that the GCE originates from a combination of dark matter and a stellar population of sub-threshold point sources (NFW $\gamma = 1.2$ & Coleman bulge). From the masks (ROIs) shown in the top panels, the backgrounds giving the best fit prefer the GCE morphology to follow the NFW $\gamma = 1.2$ & Coleman bulge choice. The two next best options are either the purely NFW $\gamma = 1.2$ profile or the Coleman bulge profile (see text for further details). The same result comes for the lower panel masks, with the difference that for the relevant background model used to get the best fit–which changes between masks–the pure dark matter annihilation profile for the GCE performs as well as the NFW $\gamma = 1.2$ & Coleman bulge profile.

NFW $\gamma = 1.2$. The other is described as the Coleman bulge profile. We are motivated to test the combined NFW $\gamma = 1.2$ & Coleman bulge based on our findings in Sec. IV C. This combination often provides the best fit to the gamma-ray data across various choices of background models and masks (see also Fig. 9).

Additionally, it is important to consider that gamma-ray point sources below the detection threshold contribute to

the data, and their influence is generally captured by other templates in our fitting process. The isotropic diffuse emission template accounts for the extragalactic sub-threshold sources. The combination of galactic diffuse emission components tends to absorb some flux from the galactic point sources, especially those correlated with the gas-dense regions of the Milky Way. By incorporating the Coleman bulge template, we explicitly test the
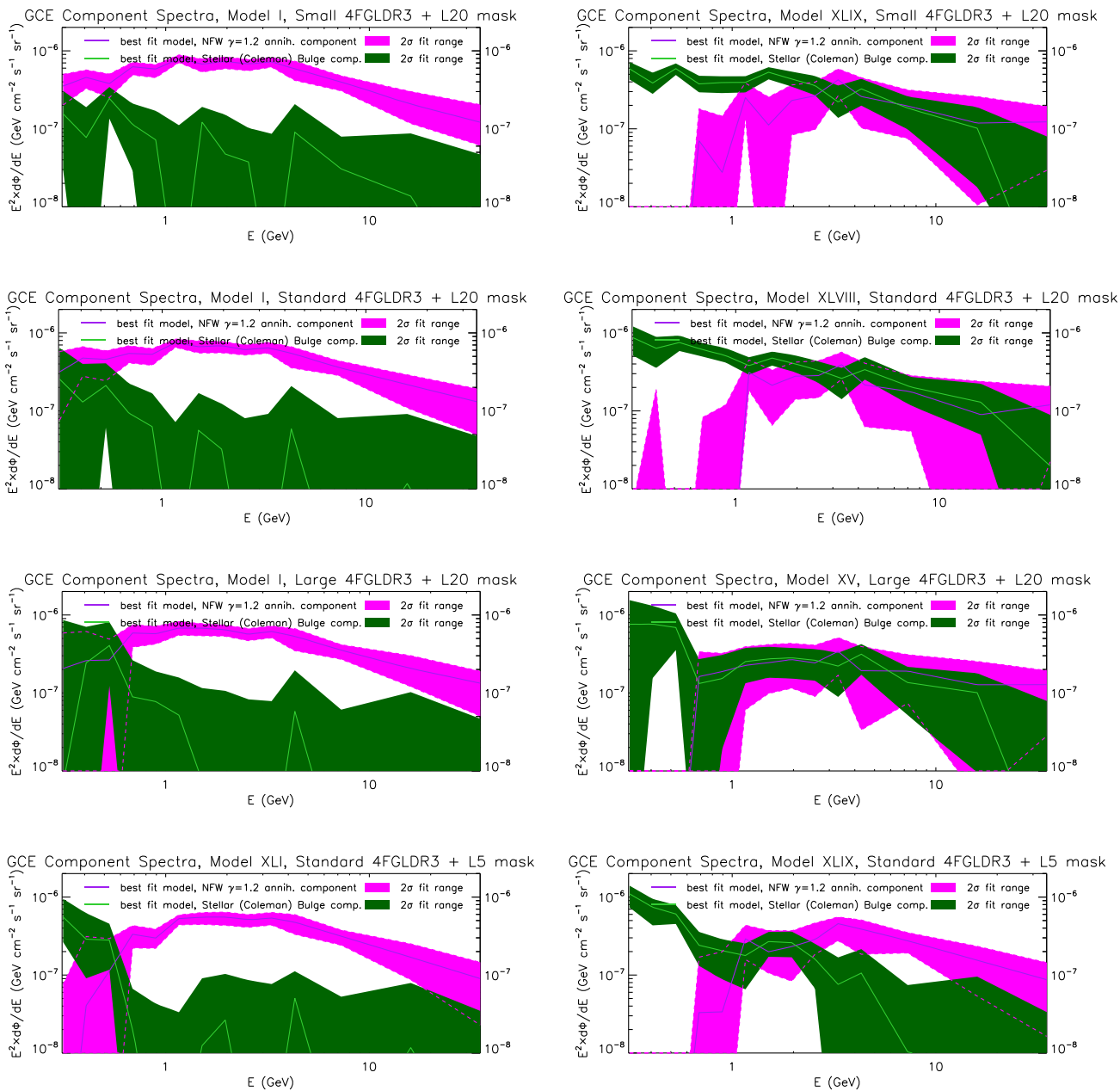
FIG. 10. Assuming that the GCE comes from the combination of dark matter and subthreshold gamma-ray point sources, we show the flux decomposition of the GCE attributed to the spherically symmetric ($\epsilon = 1.0$) dark matter annihilation template with $\gamma = 1.2$ (purple line and magenta band) and the asymmetric and elongated along the galactic disk Coleman bulge template (green line and band). We show the evolution with the energy of the two components contributing to the GCE. In each row, we show results from two different gamma-ray background models keeping the mask fixed. The two background models are selected to be among those providing the best agreement to the data. From top to bottom, we use the following masks: small 4FGLDR3 + L20, standard 4FGLDR3 + L20, large 4FGLDR3 + L20 and standard 4FGLDR3 + L5.
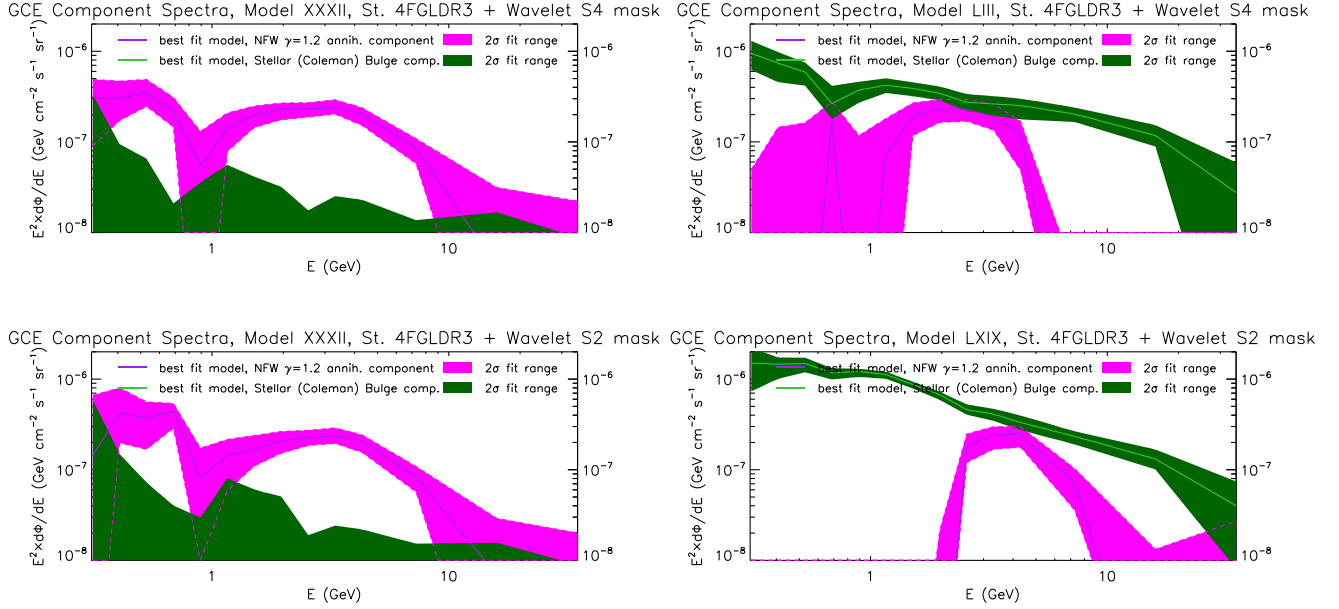
FIG. 11.   As in Fig. 10, we show the flux decomposition of the GCE attributed to the dark matter annihilation template with $\gamma = 1.2$ and $\epsilon = 1.0$ (purple line and magenta band) and the Coleman bulge template (green line and band). We show results using masks standard 4FGLDR3 $+$ wavelet S4 and standard 4FGLDR3 $+$ wavelet S2.

hypothesis that the GCE is associated morphologically with the galactic bulge.

Figures 10 and 11 show the flux decomposition of the GCE into these two components for the six different masks previously used in Figs. 6–9. Each component in our study has an independently adjustable normalization for each of the 14 energy bins. The flux attributed to the dark matter component is shown with a purple line for the best-fit normalization and a magenta band for its 2σ range, which are derived from the MCMC procedure that samples the coefficient parameters. Similarly, the flux associated with the Coleman bulge is given by a light green line for the best-fit normalization and a dark green band for its 2σ range. For each mask, we show results from two out of the 80 background models. We pick these two models among those providing a good fit to the data and in a manner that is intended to envelope for the reader the range of how much of the GCE each of the two components can absorb.

In the analysis using regular masks, as shown in Fig. 10, the dark matter component consistently comprises at least 50% of the GCE emission above 0.7 GeV, and in some instances, it accounts for over 90% of the GCE emission. Moreover, even after accounting for systematic uncertainties, the spectrum of the dark matter component is roughly the same as what has been proposed in previous works as [2–14]. This suggests that, within a factor of $\simeq 2$ adjustments in the dark matter cross section and mass, the existing explanations for the dark matter origin of the GCE remain valid, even considering all tested background diffuse emission model systematics, mask choices, and the

potential independent contribution of a galactic bulge component.

It is also notable that even when the Coleman bulge component contributes maximally (as seen in the right panels of Fig. 10), its spectrum is *distinctly different* from that typically associated with MSPs [4,8,33,58]. Instead, it more closely resembles the gamma-ray spectrum from the regular population of galactic point sources [42], which includes MSPs but also contains the emission from younger gamma-ray pulsars, including pulsar wind nebulae, sources in globular clusters, binaries, supernova remnants, and star-forming regions. This suggests that the actual contribution of regular MSPs to the GCE may be relatively minor (see, e.g., [20,22,25,32]). Results from the standard 4FGLDR1 $+$ L20, standard 4FGLDR2 $+$ L20 and standard 4FGLDR3 $+$ L8 masks, lead to the same conclusions. In Appendix B, we provide additional information on the statistical preference for either the dark matter or the Coleman bulge components of the GCE as a function of energy.

Figure 11, shows the results obtained using our wavelet-based masks standard 4FGLDR3 $+$ wavelet S2 and standard 4FGLDR3 $+$ wavelet S4. The GCE spectra derived from these masks are distinctively different from those derived when portions of the galactic disk are masked. Depending on the specific case, the dark matter component or the Coleman bulge may dominate the GCE emission. However, it is important to note that the gamma-ray spectrum associated with the Coleman bulge, again, does not resemble the known spectra of MSPs.
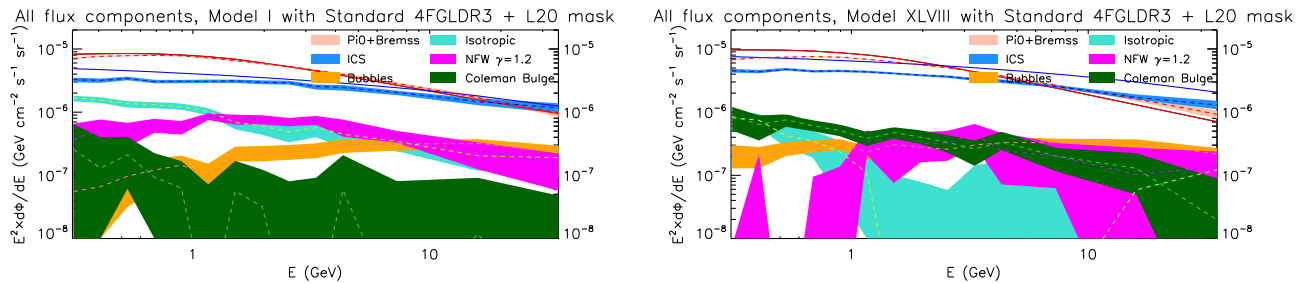
FIG. 12. For the $40° \times 40°$ ROI with the standard 4FGLDR3 + L20 mask, we provide all galactic diffuse emission components. Those are the Pi0 + Bremsstrahlung in orange lines and band, the ICS in blue lines and band, the isotropic in turquoise, *Fermi* bubbles in orange, the dark matter component in purple line and magenta band, and the Coleman bulge component in green line and band. Left panel: We use galactic diffuse emission model I of Ref. [14]. Right panel: We instead use the model XLVIII from the same reference (see text for further details).

This discrepancy suggests a significant contribution from unresolved point sources along the disk. Results from the standard 4FGLDR3 + wavelet S3 mask agree with those of the other two wavelet-based masks.

Finally, in Fig. 12, as a point of reference for the reader, we provide a complete decomposition of gamma-ray emission from the $40° \times 40°$ ROI for one of the masks and two galactic diffuse emission background models. The solid lines represent the original assumptions of the Pi0 + Bremsstrahlung and the ICS before fitting each energy bin. The dashed lines show the fitted spectra for each component, along with their corresponding 2σ ranges as colored bands. On the left panel of Fig. 12, we use the galactic diffuse emission model I of Ref. [14], where the dark matter template dominates the GCE. On the right panel, we use model XLVIII from the same reference instead, in which case the dark matter template contributes approximately half of the GCE flux above 1 GeV. Note that in the case of the right panel, the flux component associated with the Coleman bulge profile does not have the hard spectrum followed by an exponential cut-off expected by MSPs. Instead, the spectrum associated with the Coleman bulge is that expected from the population of all galactic gamma-ray point sources. Furthermore, the inferred isotropic emission component deviates from the true isotropic emission calculated at high latitudes. On the left panel, the Coleman bulge template component has a similar spectrum to that on the right but an order of magnitude dimmer. We find such spectral patterns indicative of the Coleman bulge template are likely some combination of extragalactic and galactic point sources and not an unresolved millisecond pulsar population.

V. DISCUSSIONS AND CONCLUSIONS

Masking is an essential step for the template-fitting analysis of the GCE, crucial for minimizing the influence of the galactic disk and the known point sources on the GCE emission, which are challenging to model accurately. In this study, we constructed a variety of masks based on the information of the point sources and the disk. These include

regular masks, which combine a band-shaped disk mask with point source masks based on the *Fermi* 4FGL catalogs, and wavelet-based masks, which combine masks for high-significance peaks identified by the wavelet method with masks for point sources also based on 4FGL. For the regular masks, we varied the band length for the disk mask and the radii for the point source masks.

We then extensively investigated the impact of masking on the GCE morphology analysis. Through injection tests, we demonstrated that our analysis method, the template fitting, is sensitive and robust against the ellipticity of the GCE. Through systematic analyses, we further showed that the choice of masks does not significantly impact the inferred cuspiness and ellipticity of the GCE, assuming the GCE follows contracted, squared, and integrated NFW profiles, a common template to model dark matter annihilation. This robustness in the characteristics of the GCE is evident at an energy bin-by-energy bin level, except for the lowest energy bins, due to limited statistics after masking and large astrophysical background uncertainties.

We also explored various hypotheses about the GCE's morphology based on different stellar bulge profiles, including boxy bulge, BB + NB, Coleman bulge, F98, and X-shaped bulge. We compared these results to dark matter annihilation (NFW $\gamma = 1.2$). Across different choices of masks and background diffuse emission models, we found that the NFW $\gamma = 1.2$ and the Coleman bulge profiles offer the best fits to the GCE, with the latter profile giving slightly more often a better fit to the gamma-ray data, but with a typically small difference in terms of $-2\Delta \ln \mathcal{L}$.

Additionally, we tested combinations of dark matter annihilation with the BB + NB, the F98, and the Coleman bulge. Among those combinations, we found that NFW $\gamma = 1.2$ & Coleman bulge offers the best fit across different choices of masks and background emission models. It also consistently fits better than NFW $\gamma = 1.2$ or Coleman bulge alone. This prompted further exploration of the contributions from the two components as a function of energy under different masks. We found that the dark matter

component consistently comprises a significant portion of the GCE for $E > 0.7$ GeV. Although, in some cases, the Coleman bulge can be an important component, the corresponding spectra distinctly differ from that typically associated with MSPs, resembling more the spectra of the regular galactic point sources. Besides, when the Coleman bulge is an important component, the resulting spectrum associated with the isotropic emission deviates from the measurement at high latitudes. The fact that the Coleman bulge does not yield a MSP-like spectrum is the case across both regular and wavelet-based masks. All regular masks still give a prominent spherically symmetric GCE component.

Our results using wavelet-based masks occasionally show a different GCE morphology and spectrum from those obtained with regular masks, possibly due to too much contribution from other galactic point sources. However, the template fitting based on wavelet-based masks generally maintains a consistent ranking of GCE hypotheses (dark matter annihilation, stellar bulges, and combinations) as those based on regular masks. The wavelet-based masks do accentuate the differences in $-2\Delta \ln \mathcal{L}$ between the best-fit background models and poorer-fit ones. Further exploration with a broader pool of background models may offer more suitable candidates for template fitting under wavelet-based masks.

ACKNOWLEDGMENTS

APPENDIX A: THE GCE MORPHOLOGY'S ENERGY EVOLUTION FOR ALTERNATIVE MASKS

In Sec. IV A of the main text, we presented for the standard $4\mathrm{FGLDR3} + \mathrm{L}20$ mask the evolution with energy of the GCE's cuspiness and ellipticity. Here, we show results for the other five masks, for which figures are presented in the main text. Those are the, small $4\mathrm{FGLDR3} + \mathrm{L}20$, large $4\mathrm{FGLDR3} + \mathrm{L}20$, standard $4\mathrm{FGLDR3} + \mathrm{L}5$, standard $4\mathrm{FGLDR3} +$ wavelet S4 and

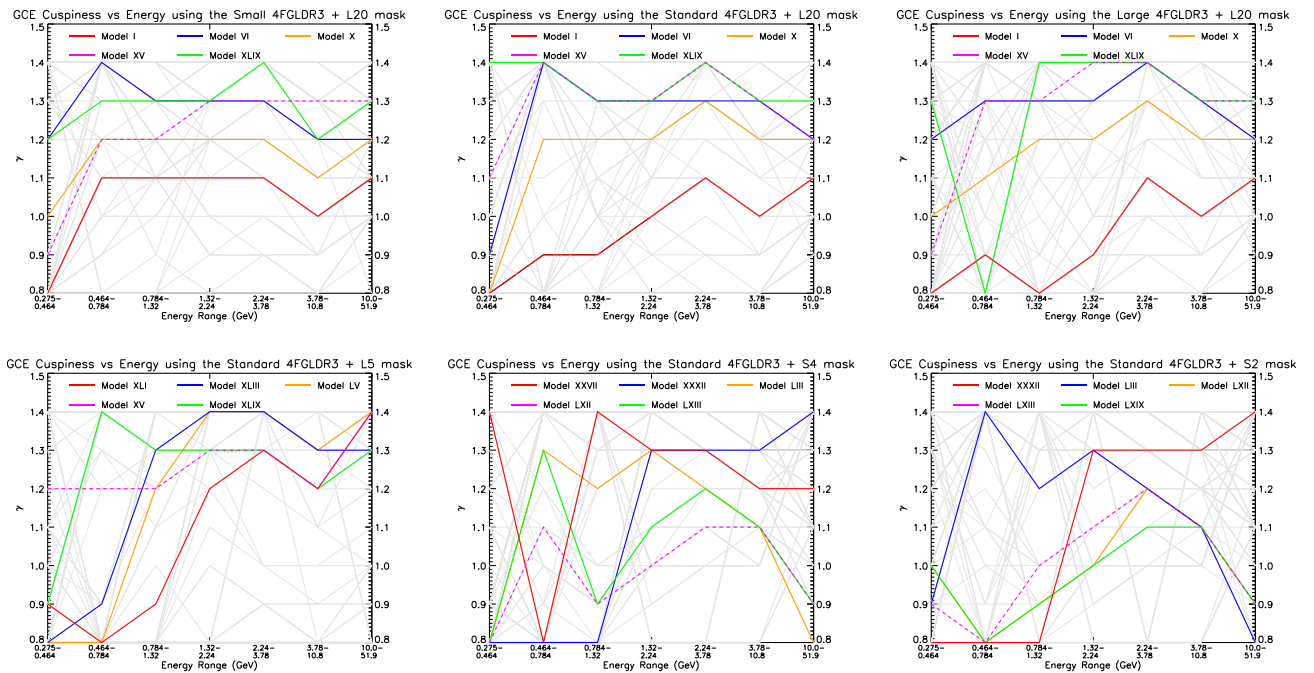

FIG. 13.   As in the left panel of Fig. 5, the evolution of the cuspiness of the GCE with energy for a sequence of masks (see text for more details).
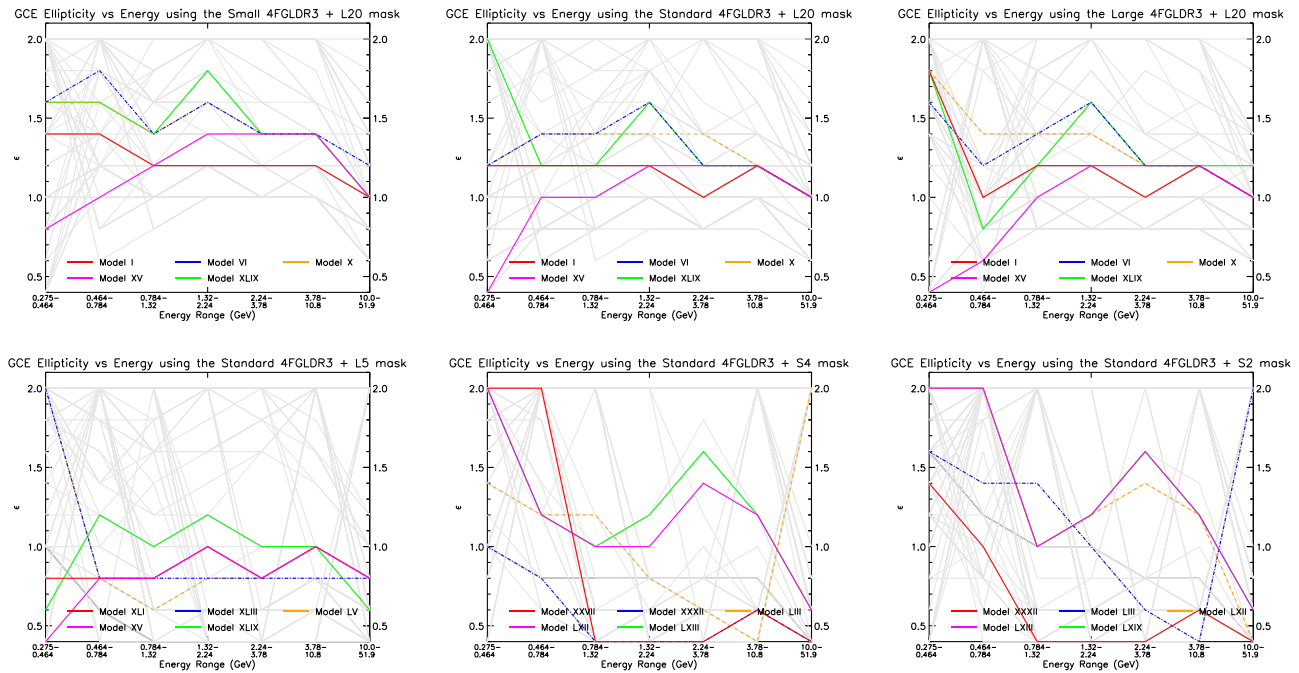
FIG. 14. As in the right panel of Fig. 5, the evolution of the ellipticity of the GCE with energy for a sequence of masks (see text for more details).

standard 4FGLDR3 + wavelet S2 masks. For ease of comparison, we also include the results of standard 4FGLDR3 + L20.

In Fig. 13, we show the cuspiness evolution with energy for these six masks. Using the regular masks (first four panels) and focusing on the background diffuse emission models that give the best fits (colored lines), the cuspiness converges to values of $1.1 \leq \gamma \leq 1.4$ for energies 3 GeV or higher. Only the wavelet-based masks provide inconclusive results on the GCE cuspiness. As described in the main text, we think that to be the result of the wavelet-based masks allowing too much contribution of other galactic (non-MSP) point sources, whose spectrum has nothing to do with the GCE emission.

In Fig. 14, we show the ellipticity evolution with energy for the same six masks. Again, the regular masks, in combination with the best-fit background models, converge to values of $0.8 \leq \epsilon \leq 1.2$ for energies > 3 GeV. Instead, the wavelet-based masks provide inconclusive results on the GCE ellipticity.

APPENDIX B: THE ENERGY EVOLUTION OF THE STATISTICAL PREFERENCE FOR THE GCE COMPONENTS

In the main text, in Secs. IV C and IV D and Figs. 9, 10 and 11 we investigated the possibility that the GCE is a combination of both a component following the dark matter annihilation profile and a component that traces the Coleman bulge profile. As we showed in the main text, the component tracing the dark matter annihilation profile is always present, with a normalization that for energies $E > 0.7$ GeV accounts for $\simeq 50$–90% of the total GCE flux. Only when we use the standard 4FGLDR3 + wavelet S2 mask, we find a lower flux associated with the dark matter annihilation profile.

In this appendix, we provide further results on the evolution with the energy of the statistical preference for the GCE to trace either the dark matter annihilation morphology or the Coleman bulge one. We test the cases where the GCE is purely tracing one or the other morphology. In Fig. 15, we show for the same six masks as in Sec. IV and for specific galactic diffuse background models the energy bin-by-energy bin difference in negative log-likelihood (in $-2\Delta \ln \mathcal{L}_j$) between the dark matter annihilation NFW with $\gamma = 1.2$ profile and the Coleman bulge profile. The biggest differences show up at energies between 1 GeV and 5 GeV, with little statistical difference between the two profiles for $E > 10$ GeV where the number of detected photons is small. Also, at $E < 0.7$ GeV, there is typically little statistical preference between the two options. At the lowest energies, the background emission is too dominant.
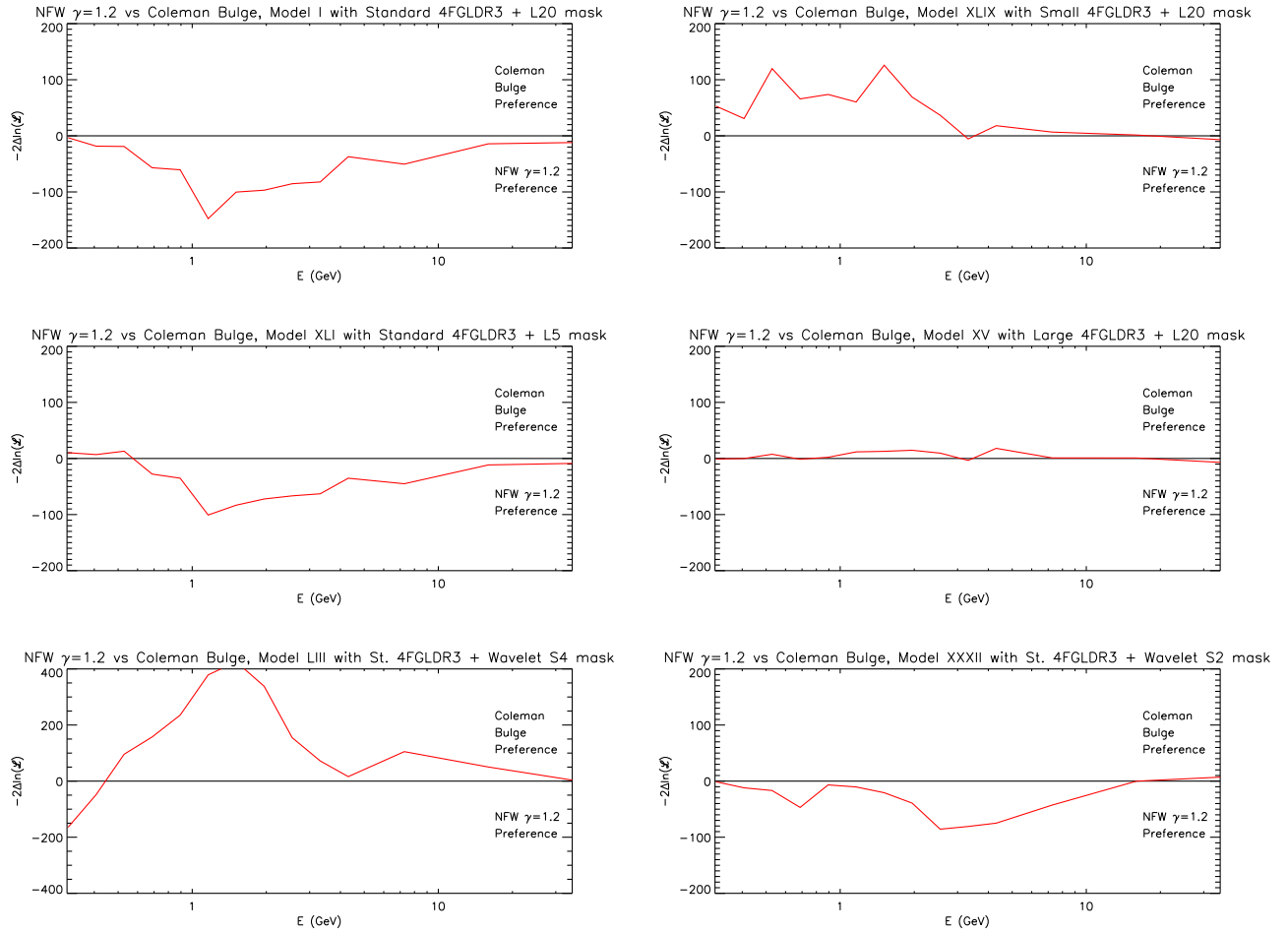
FIG. 15.   For a sample of galactic diffuse models and masks, we show the energy bin-by-energy bin level of statistical preference between the dark matter annihilation NFW with $\gamma = 1.2$ profile vs the Coleman bulge profile.

[1] Lisa Goodenough and Dan Hooper, Possible evidence for dark matter annihilation in the inner Milky Way from the Fermi gamma ray space telescope, arXiv:0910.2998.

[2] Dan Hooper and Lisa Goodenough, Dark matter annihilation in the galactic center as seen by the Fermi gamma ray space telescope, Phys. Lett. B **697,** 412 (2011).

[3] Vincenzo Vitale and Aldo Morselli, Indirect search for dark matter from the center of the Milky Way with the Fermi-Large Area Telescope, arXiv:0912.3828.

[4] Kevork N. Abazajian, The consistency of Fermi-LAT observations of the galactic center with a millisecond pulsar population in the central stellar cluster, J. Cosmol. Astropart. Phys. 03 (2011) 010.

[5] Dan Hooper and Tim Linden, On the origin of the gamma rays from the galactic center, Phys. Rev. D **84,** 123005 (2011).

[6] Dan Hooper and Tracy R. Slatyer, Two emission mechanisms in the Fermi bubbles: A possible signal of annihilating dark matter, Phys. Dark Universe **2,** 118 (2013).

[7] Chris Gordon and Oscar Macias, Dark matter and pulsar model constraints from galactic center Fermi-LAT gamma ray observations, Phys. Rev. D **88,** 083521 (2013); **89,** 049901(E) (2014).

[8] Kevork N. Abazajian, Nicolas Canac, Shunsaku Horiuchi, and Manoj Kaplinghat, Astrophysical and dark matter interpretations of extended gamma-ray emission from the galactic center, Phys. Rev. D **90,** 023526 (2014).

[9] Tansu Daylan, Douglas P. Finkbeiner, Dan Hooper, Tim Linden, Stephen K. N. Portillo, Nicholas L. Rodd, and Tracy R. Slatyer, The characterization of the gamma-ray signal from the central Milky Way: A case for annihilating dark matter, Phys. Dark Universe **12,** 1 (2016).

[10] Francesca Calore, Ilias Cholis, and Christoph Weniger, Background model systematics for the Fermi GeV excess, J. Cosmol. Astropart. Phys. 03 (2015) 038.

[11] Bei Zhou, Yun-Feng Liang, Xiaoyuan Huang, Xiang Li, Yi-Zhong Fan, Lei Feng, and Jin Chang, GeV excess in the Milky Way: The role of diffuse galactic gamma-ray emission templates, Phys. Rev. D **91,** 123010 (2015).

[12] M. Ajello *et al.* (Fermi-LAT Collaboration), Fermi-LAT observations of high-energy γ-ray emission toward the galactic center, Astrophys. J. **819,** 44 (2016).

[13] Tim Linden, Nicholas L. Rodd, Benjamin R. Safdi, and Tracy R. Slatyer, High-energy tail of the galactic center gamma-ray excess, Phys. Rev. D **94,** 103013 (2016).

[14] Ilias Cholis, Yi-Ming Zhong, Samuel D. McDermott, and Joseph P. Surdutovich, Return of the templates: Revisiting the galactic center excess with multimessenger observations, Phys. Rev. D **105,** 103023 (2022).

[15] Francesca Calore, Ilias Cholis, Christopher McCabe, and Christoph Weniger, A tale of tails: Dark matter interpretations of the Fermi GeV excess in light of background model systematics, Phys. Rev. D **91,** 063003 (2015).

[16] Prateek Agrawal, Brian Batell, Patrick J. Fox, and Roni Harnik, WIMPs at the galactic center, J. Cosmol. Astropart. Phys. 05 (2015) 011.

[17] Asher Berlin, Stefania Gori, Tongyan Lin, and Lian-Tao Wang, Pseudoscalar portal dark matter, Phys. Rev. D **92,** 015005 (2015).

[18] Christopher Karwin, Simona Murgia, Tim M. P. Tait, Troy A. Porter, and Philip Tanedo, Dark matter interpretation of the Fermi-LAT observation toward the galactic center, Phys. Rev. D **95,** 103005 (2017).

[19] M. Ackermann *et al.* (Fermi-LAT Collaboration), The Fermi galactic center GeV excess and implications for dark matter, Astrophys. J. **840,** 43 (2017).

[20] Rebecca K. Leane and Tracy R. Slatyer, Revival of the dark matter hypothesis for the galactic center gamma-ray excess, Phys. Rev. Lett. **123,** 241101 (2019).

[21] Kevork N. Abazajian and Manoj Kaplinghat, Detection of a gamma-ray source in the galactic center consistent with extended emission from dark matter annihilation and concentrated astrophysical emission, Phys. Rev. D **86,** 083511 (2012); **87,** 129902(E) (2013).

[22] Dan Hooper, Ilias Cholis, Tim Linden, Jennifer Siegal-Gaskins, and Tracy Slatyer, Pulsars cannot account for the inner galaxy's GeV excess, Phys. Rev. D **88,** 083009 (2013).

[23] Samuel K. Lee, Mariangela Lisanti, Benjamin R. Safdi, Tracy R. Slatyer, and Wei Xue, Evidence for unresolved γ-ray point sources in the inner galaxy, Phys. Rev. Lett. **116,** 051103 (2016).

[24] Jovana Petrović, Pasquale D. Serpico, and Gabrijela Zaharijas, Millisecond pulsars and the galactic center gamma-ray excess: The importance of luminosity function and secondary emission, J. Cosmol. Astropart. Phys. 02 (2015) 023.

[25] Ilias Cholis, Dan Hooper, and Tim Linden, Challenges in explaining the galactic center gamma-ray excess with millisecond pulsars, J. Cosmol. Astropart. Phys. 06 (2015) 043.

[26] Richard Bartels, Emma Storm, Christoph Weniger, and Francesca Calore, The Fermi-LAT GeV excess as a tracer

[27] Malte Buschmann, Nicholas L. Rodd, Benjamin R. Safdi, Laura J. Chang, Siddharth Mishra-Sharma, Mariangela Lisanti, and Oscar Macias, Foreground mismodeling and the point source explanation of the Fermi galactic center excess, Phys. Rev. D **102,** 023023 (2020).

[28] Anuj Gautam, Roland M. Crocker, Lilia Ferrario, Ashley J. Ruiter, Harrison Ploeg, Chris Gordon, and Oscar Macias, Millisecond pulsars from accretion-induced collapse as the origin of the galactic centre gamma-ray excess signal, Nat. Astron. **6,** 703 (2022).

[29] Jovana Petrović, Pasquale Dario Serpico, and Gabrijela Zaharijaš, Galactic center gamma-ray "excess" from an active past of the galactic centre?, J. Cosmol. Astropart. Phys. 10 (2014) 052.

[30] Eric Carlson and Stefano Profumo, Cosmic ray protons in the inner galaxy and the galactic center gamma-ray excess, Phys. Rev. D **90,** 023015 (2014).

[31] Ilias Cholis, Carmelo Evoli, Francesca Calore, Tim Linden, Christoph Weniger, and Dan Hooper, The galactic center GeV excess from a series of leptonic cosmic-ray outbursts, J. Cosmol. Astropart. Phys. 12 (2015) 005.

[32] Yi-Ming Zhong, Samuel D. McDermott, Ilias Cholis, and Patrick J. Fox, Testing the sensitivity of the galactic center excess to the point source mask, Phys. Rev. Lett. **124,** 231103 (2020).

[33] Richard Bartels, Suraj Krishnamurthy, and Christoph Weniger, Strong support for the millisecond pulsar origin of the galactic center GeV excess, Phys. Rev. Lett. **116,** 051102 (2016).

[34] Oscar Macias, Chris Gordon, Roland M. Crocker, Brendan Coleman, Dylan Paterson, Shunsaku Horiuchi, and Martin Pohl, Galactic bulge preferred over dark matter for the galactic centre gamma-ray excess, Nat. Astron. **2,** 387 (2018).

[35] Oscar Macias, Shunsaku Horiuchi, Manoj Kaplinghat, Chris Gordon, Roland M. Crocker, and David M. Nataf, Strong evidence that the galactic bulge is shining in gamma rays, J. Cosmol. Astropart. Phys. 09 (2019) 042.

[36] Mattia Di Mauro, Characteristics of the galactic center excess measured with 11 years of *Fermi*-LAT data, Phys. Rev. D **103,** 063029 (2021).

[37] Martin Pohl, Oscar Macias, Phaedra Coleman, and Chris Gordon, Assessing the impact of hydrogen absorption on the characteristics of the galactic center excess, Astrophys. J. **929,** 136 (2022).

[38] Samuel D. McDermott, Yi-Ming Zhong, and Ilias Cholis, On the morphology of the gamma-ray galactic centre excess, Mon. Not. R. Astron. Soc. **522,** L21 (2023).

[39] Andrew L. Miller and Yue Zhao, Probing the pulsar explanation of the galactic-center GeV excess using continuous gravitational-wave searches, Phys. Rev. Lett. **131,** 081401 (2023).

[40] S. Abdollahi *et al.* (Fermi-LAT Collaboration), *Fermi* large area telescope fourth source catalog, Astrophys. J. Suppl. Ser. **247,** 33 (2020).

[41] J. Ballet, T. H. Burnett, S. W. Digel, and B. Lott (Fermi-LAT Collaboration), Fermi Large Area Telescope fourth source catalog data release 2, arXiv:2005.11208.

[42] Soheila Abdollahi *et al.* (Fermi-LAT Collaboration), Incremental Fermi Large Area Telescope fourth source catalog, Astrophys. J. Suppl. Ser. **260,** 53 (2022).

[43] J. Ballet, P. Bruel, T. H. Burnett, and B. Lott (Fermi-LAT Collaboration), Fermi Large Area Telescope fourth source catalog data release 4 (4FGL-DR4), Astrophys. J. Suppl. Ser. **260,** 53 (2022).

[44] Brenna Coleman, Dylan Paterson, Chris Gordon, Oscar Macias, and Harrison Ploeg, Maximum entropy estimation of the galactic bulge morphology via the VVV red clump, Mon. Not. R. Astron. Soc. **495,** 3350 (2020).

[45] Fermi-LAT performance (2021), https://www.slac.stanford .edu/exp/glast/groups/canda/lat_Performance.htm.

[46] M. Ackermann *et al.* (Fermi-LAT Collaboration), The spectrum and morphology of the *Fermi* bubbles, Astrophys. J. **793,** 64 (2014).

[47] M. Ackermann *et al.* (Fermi-LAT Collaboration), The spectrum of isotropic diffuse gamma-ray emission between 100 MeV and 820 GeV, Astrophys. J. **799,** 86 (2015).

[48] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman, emcee: The MCMC Hammer, Publ. Astron. Soc. Pac. **125,** 306 (2013).

[49] S. R. Hinton, CHAINCONSUMER, J. Open Source Software **1,** 45 (2016).

[50] Robert J. J. Grand and Simon D. M. White, Dark matter annihilation and the galactic centre excess, Mon. Not. R. Astron. Soc. **511,** L55 (2022).

[51] J. Diemand, M. Kuhlen, P. Madau, M. Zemp, B. Moore, D. Potter, and J. Stadel, Clumps and streams in the local dark matter distribution, Nature (London) **454,** 735 (2008).

[52] M. Kuhlen, J. Diemand, P. Madau, and M. Zemp, The via lactea INCITE simulation: Galactic dark matter substructure at high resolution, J. Phys. Conf. Ser. **125,** 012008 (2008).

[53] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White, The aquarius project: The subhaloes of galactic haloes, Mon. Not. R. Astron. Soc. **391,** 1685 (2008).

[54] Gregory Dobler, Ilias Cholis, and Neal Weiner, The Fermi gamma-ray haze from dark matter annihilations and anisotropic diffusion, Astrophys. J. **741,** 25 (2011).

[55] R. Launhardt, R. Zylka, and P. G. Mezger, The nuclear bulge of the Galaxy. III. Large-scale physical characteristics of stars and interstellar matter, Astron. Astrophys. **384,** 112 (2002).

[56] Oscar Macias, Deheng Song, and Manoj Kaplinghat (private communication).

[57] H. T. Freudenreich, Cobe's galactic bar and disk, Astrophys. J. **492,** 495 (1998).

[58] Ilias Cholis, Dan Hooper, and Tim Linden, A new determination of the spectra and luminosity function of gamma-ray millisecond pulsars, arXiv:1407.5583.

[59] A. W. Strong and I. V. Moskalenko, Propagation of cosmic-ray nucleons in the Galaxy, Astrophys. J. **509,** 212 (1998).