Search for dark matter annihilation in the Milky Way halo

Laura J. Chang, Mariangela Lisanti, and Siddharth Mishra-Sharma

A Search for Dark Matter Annihilation in the Milky Way Halo

Laura J. Chang,[1] Mariangela Lisanti,[1] and Siddharth Mishra-Sharma[1]

[1]*Department of Physics, Princeton University, Princeton, NJ 08544*

(Dated: September 14, 2018)

The Milky Way halo is the brightest source of dark matter annihilation on the sky. Indeed, the potential strength of the Galactic dark matter signal can supersede that expected from dwarf galaxies and galaxy groups even in regions away from the Inner Galaxy. In this paper, we present the results of a search for dark matter annihilation in the smooth Milky Way halo for $|b| > 20°$ and $r < 50°$ using 413 weeks of *Fermi* Pass 8 data within the energy range of $\sim$0.8–50 GeV. We exclude thermal dark matter with mass below $\sim$70 GeV that annihilates to $b\bar{b}$ at the 95% confidence level using the p6v11 cosmic-ray foreground model, providing the strongest limits on the annihilation cross section in this mass range. These results exclude the region of dark matter parameter space that is consistent with the excess of $\sim$GeV photons observed at the Galactic Center for the $b\bar{b}$ annihilation channel and, for the first time, start probing the $\tau^+\tau^-$ explanation. We explore how these results depend on uncertainties in the foregrounds by varying over a set of reasonable models.

I. INTRODUCTION

The *Fermi* Large Area Telescope [1] provides an unprecedented view of the gamma-ray sky. The all-sky maps that are available can harbor clues about the nature of dark matter (DM), which can annihilate to visible states that produce showers of high-energy photons. A variety of such searches have been performed, focusing on regions where the relative DM density is expected to be significant. Thus far, the most sensitive bounds come from looking at ultrafaint dwarf galaxies [2–5] and galaxy groups [6, 7]. In this paper, we explore emission due to annihilating DM from the Galactic halo, and demonstrate that it can be used to set robust constraints on the DM annihilation cross section. These constraints are the strongest to date on DM with mass less than $\sim$70 GeV, for the $b\bar{b}$ annihilation benchmark.

The halo surrounding our Galaxy provides the brightest source of DM emission on the sky. In general, the DM flux is proportional to the so-called J-factor, which is the integral over the line-of-sight, s, and solid angle, Ω, of the squared DM density profile:

$$J = \int ds \, d\Omega \, \rho^2(s, \Omega). \tag{1}$$

The J-factor provides a useful metric for comparing the strength of an annihilation signal expected from different targets. For example, the J-factors from some of the brightest ultrafaint dwarf galaxies are $\sim$10^{19} GeV^2 cm^{-5} sr [5], comparable to those of the brightest galaxy groups [6]. In contrast, the center of our own Galaxy has a J-factor several orders of magnitude larger, with $J \sim 10^{23}$ GeV^2 cm^{-5} sr in the inner $40° \times 40°$ region. Even if one were to avoid the central part of the Galaxy and only consider an annulus of $r < 50°$ and latitudes greater than $|b| > 20°$, the J-factor is still as large as $\sim$10^{22} GeV^2 cm^{-5} sr.

Despite the strength of the smooth Galactic DM signal, many other factors complicate a potential search. The primary challenge is posed by the bright diffuse emission from cosmic rays propagating in the Galaxy. These contributions arise from π^0 decay, Bremsstrahlung from the interaction of cosmic-ray electrons with interstellar gas, and inverse-Compton (IC) scattering of photons off of high-energy electrons. This diffuse foreground contributes the vast majority of the high-energy photons we see on the sky, accounting for $\sim$50–90% of the observed photons depending on the energy range considered [8], and is challenging to model accurately. Any search for Galactic DM must mitigate these uncertainties and quantify the effects of varying over assumptions in the foreground models.

Searches for Galactic DM can be divided into two broad categories. The first set focuses on the Inner Galaxy, within $r \lesssim 20°$ [9–24]. These analyses have conclusively found an excess of $\sim$GeV photons whose energy distribution and spatial morphology can be consistent with the expectation due to DM annihilation. However, recent studies have shown that the distribution of photons in the Inner Galaxy is more consistent with a population of unresolved point sources, disfavoring the DM interpretation [25, 26]. Additionally, other studies suggest that the spatial morphology of the excess may better trace the stellar bulge [27, 28]. Complementary studies of Milky Way dwarfs [5] and galaxy groups [6, 7] are starting to put in tension the DM interpretation of the excess emission. However, the tension can be alleviated depending on the specific assumptions made about, *e.g.*, the dwarf halo profiles [3, 29, 30]; the stellar membership criteria used to infer the dwarf halo properties [31, 32]; the shape of the Milky Way halo [33]; or the nature of substructure boost in galaxy groups [7].

Even though the Galactic Center is the brightest DM source on the sky, it is also one of the most complicated due to the large astrophysical foregrounds. A complementary approach to looking at the Inner Galaxy relies on looking at the Galactic halo at higher latitudes where the DM density is still large, but the foreground levels are much smaller [34–40]. This is the approach that we take in this work. Focusing on a region defined by $|b| > 20°$ and $r < 50°$, we search for signals of DM annihilation from the smooth Milky Way halo. The limits obtained

provide the strongest constraints on low-mass DM annihilation signals and tightly constrain the DM interpretation of the GeV Excess. We verify the robustness of these results in the presence of a potential DM signal and discuss how they are affected by variations in the Galactic foreground models.

This paper is organized as follows. In Sec. II we describe our analysis pipeline and the statistical procedure employed. In Sec. III we present the results of our study, discussing the effects of Galactic foreground mismodeling and steps taken to reduce their impact. We also discuss the implications of our results for the DM interpretation of the Galactic Center Excess. We conclude in Sec. IV.

II. ANALYSIS PROCEDURE

We make use of 413 weeks of *Fermi*-LAT Pass 8 data collected between August 4, 2008 and July 7, 2016. We analyze the subset of photons in the `ULTRACLEANVETO` event class, restricting to the top quarter of photons by quality of point-spread function (PSF) reconstruction (corresponding to PSF3 event type). The data is binned in 18 logarithmically-spaced energy bins between $\sim$0.8–50 GeV. The recommended quality cuts are applied, corresponding to zenith angle less than $90°$, `LAT_CONFIG = 1`, and `DATA_QUAL > 0`.[1] Each energy bin is spatially binned into individual pixels using `HEALPix` [41] with `nside = 128`; the dataset is thus reduced to an array of integers that describes the number of photons in the energy bin, i, and pixel, p.

Template fitting is a standard astrophysical procedure where the data is described by a set of spatial maps (referred to as templates) that are binned in the same way as the data, which describe the separate components that contribute to the total photon count. Each template is associated with a normalization that is treated as a free model parameter in the fit. The likelihood for a given energy bin is then a product of the Poisson probabilities associated with the observed counts n_i^p in each pixel of the region-of-interest:

$$\mathcal{L}_i(d_i|\boldsymbol{\theta}_i) = \prod_p \frac{\mu_i^p(\boldsymbol{\theta}_i)^{n_i^p} e^{-\mu_i^p(\boldsymbol{\theta}_i)}}{n_i^p!} , \qquad (2)$$

where d_i denotes the data in energy bin i, $\boldsymbol{\theta}_i$ represents the set of model parameters and $\mu_i^p(\boldsymbol{\theta}_i)$ is the number of expected counts in a given pixel and energy bin. The total likelihood is simply the product over the individual $\mathcal{L}_i$ for each energy bin.

The region-of-interest (ROI) for this study is chosen to maximize the strength of the DM signal while minimizing the effects of foreground mis-modeling. Specifically, we take $|b| > 20°$ to avoid the Galactic plane, where cosmic-ray emission is particularly bright and there are more unresolved point sources. In addition, we take $r < 50°$ to reduce the possibility of over-subtraction and/or spurious excesses obtained from fitting the foreground model over large sky areas. Better modeling of the Galactic diffuse emission is an ongoing effort [42–46], and the use of image reconstruction and parametric modeling techniques such as SKYFACT [47] and D³PO [40, 48] could improve modeling of the Galactic diffuse emission over significantly larger regions of the sky. Increasing the radial cut in r beyond that used here could potentially improve sensitivities to DM by $\sim$20–30% or more, depending on the density profile—see App. A for a discussion of Asimov projections.

The expected photon count, $\mu_i^p(\boldsymbol{\theta}_i)$, in each pixel of the ROI depends on contributions from standard astrophysical sources as well as DM, if present. We account for four astrophysical components that trace: (i) the Galactic diffuse emission, as described by the *Fermi* `gll_iem_v02_P6_V11_DIFFUSE` (p6v11) model,[2] (ii) the *Fermi* bubbles [49], (iii) isotropic emission, and (iv) *Fermi* 3FGL point sources [50]. The smooth Galactic DM template is modeled using a generalized Navarro-Frenk-White (NFW) profile [51]:

$$\rho_{\text{NFW}}(r) = \frac{\rho_0}{(r/r_s)^\gamma [1 + (r/r_s)]^{3-\gamma}} \qquad (3)$$

with inner-slope $\gamma = 1$, scale radius $r_s = 17$ kpc, and local density $\rho(r_\odot) = 0.4$ GeV cm^{-3} [52, 53] at the Solar position $r_\odot = 8$ kpc [54]. All templates are smoothed with the energy-dependent PSF of the LAT instrument, modeled as a King function.[3]

In our fiducial study, we use the p6v11 Galactic diffuse emission model, which is designed to capture changes in the cosmic-ray emission on the full sky as a function of Galactocentric radius. The model includes contributions from π^0-decay and Bremsstrahlung emission, as traced by maps of gas column-densities, as well as inverse-Compton emission, as predicted using GALPROP [55]; the relative normalizations of these separate components are fixed. The fact that the p6v11 model should be used with caution for energies above $\sim$50 GeV sets our upper energy cut-off. To give the p6v11 template more freedom, we divide it into eight radial slices of equal area within our ROI. The normalization of each slice is then varied separately in the fitting procedure. Each slice is roughly $\sim$440 deg^2 in area, comparable in size to the regions used in dwarf and galaxy group studies ($\sim$100 and 316 deg^2, respectively) [5, 7] and smaller than the typical regions used in Inner Galaxy analyses ($\sim$1600 deg^2) [19, 21]. The

---

[1] https://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/Cicerone_Data_Exploration/Data_preparation.html

[2] https://fermi.gsfc.nasa.gov/ssc/data/access/lat/ring_for_FSSC_final4.pdf

[3] https://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/Cicerone_LAT_IRFs/IRF_PSF.html

additional freedom given to the Galactic diffuse template allows the fit to better account for localized excesses or mis-modeled features in the emission. Note that we do not use the `gll_iem_v06` (p8R2) model [56], which is recommended for the Pass 8 dataset. `p8R2` includes large-scale residuals obtained from a fit to the *Fermi* data that have been added back into the model, and is therefore not appropriate when searching for extended DM signals. We also do not use the `p7v6` diffuse model, which contains large-scale structures including Loop I and the *Fermi* bubbles with a fixed normalization.

The `p6v11` model does not include known large-scale structures that overlap with the ROI, such as the *Fermi* bubbles. We account for the bubbles by adding two templates that model the Northern and Southern lobes. The shape of the lobes is inferred directly from *Fermi* data [49], and the intensity of the emission is taken to be flat. We let the normalization of the Northern and Southern lobes float independently in the fit. The ROI also overlaps with Loop I, a large radio lobe in the Northern hemisphere [57, 58]. While features corresponding to the radio observations have been observed in the *Fermi* data [49, 59–61], significant uncertainties remain in the modeling of the spatial and intensity profile of Loop I in gamma rays. As a result, we conservatively do not include a template that traces Loop I in our fiducial study. We have performed variants of the fiducial study to assess the impact of this choice. We find that the inclusion of an additional isotropic template in the Northern hemisphere as a proxy for Loop I emission strengthens the limit by a factor of $\lesssim 1.2$.

The isotropic template is intended to primarily capture extragalactic gamma-ray emission from unresolved sources such as blazars and star-forming galaxies, as well as more exotic contributions from extragalactic DM annihilation. The inclusion of the point-source template accounts for emission from resolved (Galactic and extragalactic) sources. The normalizations of all the sources are floated together in the template after fixing their individual fluxes to the values predicted by the 3FGL catalog. We note that all 3FGL sources are conservatively masked to 95% containment in PSF for the corresponding energy bins. Therefore, the primary purpose of the point-source template is to account for any potential mis-modeling in the tails of the emission.

To summarize, there are twelve free parameters associated with the astrophysical components—eight for the Galactic diffuse slices, two for the *Fermi* bubbles, and one each for the isotropic and point-source templates. As we are ultimately interested in the intensity of the DM signal, we treat these as nuisance parameters and remove them using the profile likelihood method [62]. Specifically, we build a likelihood profile for the intensity associated with DM annihilation in the smooth Galactic halo, fixing the normalization of this template at various values while profiling over the astrophysical components. The resulting likelihood only depends on the DM intensity in each energy bin, which is related to the annihilation cross

section, $\langle \sigma v \rangle$, and mass, m_χ, through the expression for the differential gamma-ray flux:

$$\frac{d\Phi}{dE_\gamma} = J \times \frac{\langle \sigma v \rangle}{8\pi m_\chi^2} \sum_j \mathrm{Br}_j \frac{dN_j}{dE_\gamma}, \qquad (4)$$

where E_γ is the gamma-ray energy and Br_j is the branching fraction to the jth annihilation channel. The energy spectrum for each channel is described by the function dN_j/dE_γ, which is modeled using `PPPC4DMID` [63]. Note that we do not account for DM substructure in the Milky Way halo in this study, which would increase the strength of the annihilation signal. Given the theoretical uncertainties associated with modeling the spatial distribution and properties of DM subhalos, such a search deserves its own dedicated study.

The test statistic (TS) profile for $\langle \sigma v \rangle$ is defined as

$$\mathrm{TS} \equiv 2 \left[\log \mathcal{L}(d|\mathcal{M}, \langle \sigma v \rangle, m_\chi) - \log \mathcal{L}(d|\mathcal{M}, \widehat{\langle \sigma v \rangle}, m_\chi) \right], \tag{5}$$

where $\widehat{\langle \sigma v \rangle}$ is the cross section that maximizes the likelihood for a specified DM model, $\mathcal{M}$, of given annihilation channel and mass. The TS is nonpositive by definition and can be used to set a threshold for limits on the cross section. In particular, the 95% upper limit on the annihilation cross section is given by the value of $\langle \sigma v \rangle$ associated with $\mathrm{TS} = -2.71$. We implement template fitting with the package `NPTFit` [64] and use the L-BFGS-B [65] minimization algorithm implemented through `SciPy` [66].

We have performed numerous tests to ensure that the statistical procedure outlined above can recover a potential signal in the data. Such tests are crucial in verifying the robustness of these methods, especially given the potentially large degeneracies between the signal and foreground components, which are both diffuse in nature. Additionally, the freedom given to the foreground emission by separately fitting its normalization in the radial slices can lead to challenges in regimes of low photon statistics. We have performed tests on both data and Monte Carlo and verified that our analysis procedure would not exclude a DM signal if one were present in the data. A detailed description of these tests is provided in App. B.

III. RESULTS AND DISCUSSION

A. Dark Matter Annihilation Limit

Figure 1 shows the 95% confidence limit on the DM annihilation cross section into the $b\bar{b}$ final state (solid black). For comparison, the published limits from the most recent dwarf [5] and galaxy group [7] studies appear as the grey dashed and dot-dashed lines, respectively. The $b\bar{b}$ limits from the smooth Galactic halo are the strongest to date for DM masses below ~ 70 GeV.
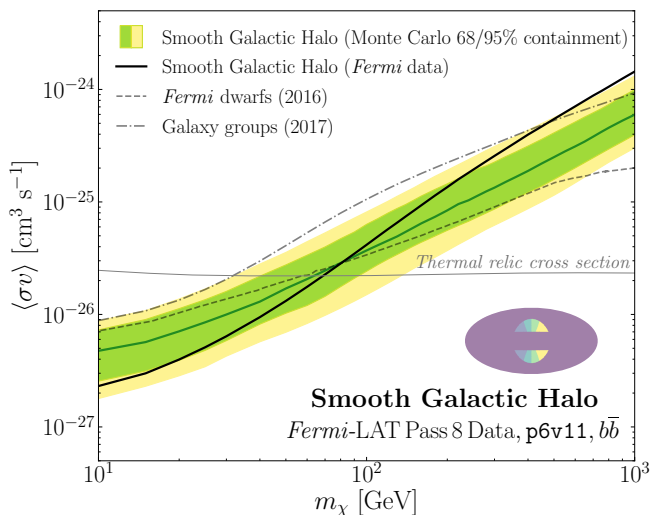
FIG. 1. The solid black line shows the 95% confidence limit on dark matter of mass, m_χ, annihilating with cross section, $\langle\sigma v\rangle$, in the smooth Galactic halo, within $|b| > 20°$ and $r < 50°$, obtained using the p6v11 foreground model. The green(yellow) band denotes the 68(95)% containment for the expected sensitivity, as derived from Monte Carlo simulations. For the Galactic halo, we assume a generalized NFW profile with inner slope of $\gamma = 1$ and local density $\rho(r_\odot) = 0.4$ GeV cm^{-3}. We also show the corresponding limits obtained from dwarf galaxies [5] and galaxy groups [7] (grey dashed and dot-dashed lines, respectively). The expected annihilation cross section for a generic weakly interacting massive particle is indicated by the solid grey line [67]. The inset depicts the eight radially sliced regions within the fiducial ROI over which the p6v11 template is allowed to float.

The green(yellow) band in Fig. 1 shows the 68(95)% expected sensitivity obtained from Monte Carlo simulations. To make the simulations, we Poisson fluctuate the sum of best-fit templates on data within the ROI, letting the normalizations for the different foreground slices and bubble lobes float independently. The sensitivity projection is derived from 100 Monte Carlo variations. A data-driven foreground expectation obtained by looking at a large number of blank fields, as is standard for dwarf and galaxy group studies, is not feasible for Galactic DM searches because the overall size of the ROI is a substantial fraction of the full sky. The Monte Carlo bands do, however, provide an important comparison benchmark. For example, if the Galactic foregrounds are over-subtracted in the fitting procedure, then the data limits will be artificially strengthened and appear stronger than the Monte Carlo expectation.

While the morphology of the signal template suggests that one should minimize the latitude cut ($|b| > b_{\rm cut}$) and maximize the radial cut ($r < r_{\rm cut}$) for optimal sensitivity to DM (see App. A for more details), a full-sky analysis is not viable in actuality due to the large uncertainties associated with modeling the Galactic foregrounds. As a result, we conservatively choose $b_{\rm cut} = 20°$

to avoid the Galactic plane, where the foregrounds are particularly bright and there is increased contamination from unresolved point sources. In addition, we choose $r_{\rm cut} = 50°$ because fitting over larger sky regions can lead to over-subtraction and/or spurious excesses in the data analysis. While the definition of the fiducial ROI is intended to mitigate the large systematic uncertainties associated with the foregrounds, we also give the p6v11 template additional freedom by fitting its normalization separately in eight radial slices of equal area, as discussed in Sec. II. Figure 2 demonstrates the need for these additional steps. The left panel shows the data limit and corresponding Monte Carlo expectation obtained when the p6v11 template is not divided into eighths, for our fiducial ROI. The right panel shows the case corresponding to a larger radial cut $r_{\rm cut} = 100°$. Every other aspect of the analysis is kept the same as in the fiducial study in these cases, except that the Northern and Southern lobes of the *Fermi* bubbles are floated together.[4] The projected sensitivities obtained from Monte Carlo simulations are essentially equivalent between the fiducial study and these two examples. The data limits, on the other hand, are starkly different. A large excess in the data limit compared to the Monte Carlo expectation is apparent when the larger ROI is used. When the fiducal ROI is used but the foreground template is not broken into radial slices, over-subtraction leads to artificially strong bounds. We therefore conclude that performing the fit over smaller sky regions and varying the p6v11 template over additional degrees of freedom stabilizes the analysis in the designated ROI.

B. Galactic Foreground Modeling

Uncertainties due to modeling of the Galactic diffuse emission are inherent in searches for large-scale gamma-ray structures. We have made an effort to minimize the effects of these uncertainties by giving more degrees of freedom to the p6v11 template. However, inherent assumptions that go into the construction of the template can still have a potentially large effect on the final result. Here, we present results for three additional foreground models that are designed to span several well-motivated possibilities. Our approach is to understand how each set of assumptions regarding the cosmic-ray modeling impacts the DM sensitivity for the ROI considered in this work.

We repeat the analysis using Models A, B, and C, which were developed by the *Fermi*-LAT Collaboration specifically for their study of the isotropic gamma-ray background at higher latitudes [8]. These models make distinct but well-motivated choices for the cosmic-ray source distribution, diffusion coefficients, and re-

---

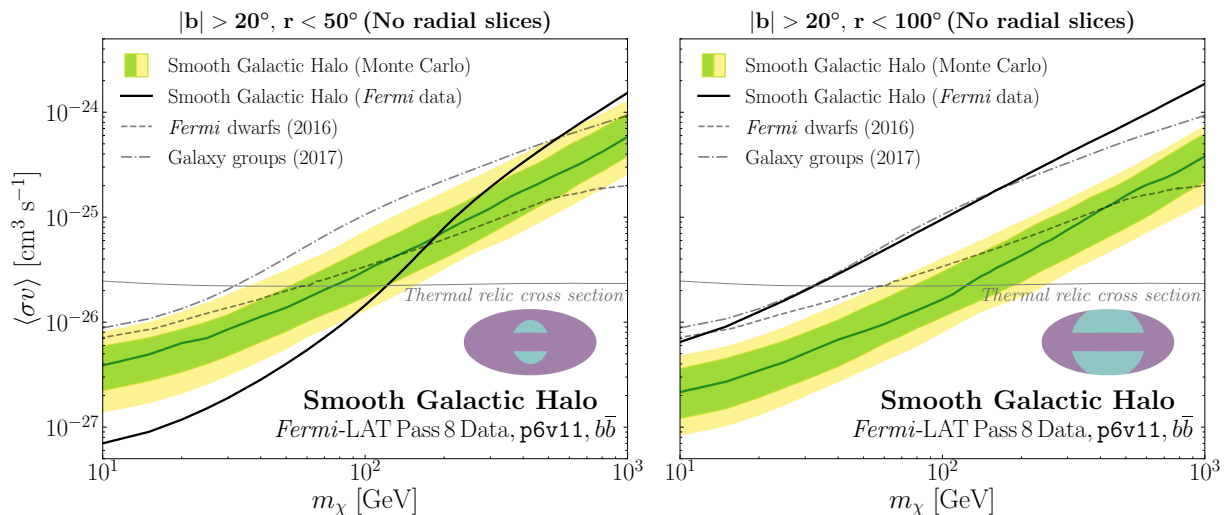[4] Doing the same for the fiducial study does not change the result.

FIG. 2. The limit when the `p6v11` template is not divided into eight radial slices whose normalizations float independently in the fitting procedure. For each panel, the inset depicts the regions (denoted in turquoise) over which the `p6v11` template is allowed to float. A region corresponding to $|b| > 20°$ and $r < 50°(100°)$ is used in the left(right) panel.

acceleration strengths that span a wide range of possibilities. Separate templates for π^0 decay, Bremsstrahlung, and inverse-Compton (IC) emission are provided, so their normalizations can be varied independently in the fitting procedure. In these analyses, we use a single combination of the Bremsstrahlung and π^0-decay templates as obtained from a fit to data using eight separate equal-area slices. Both these components trace the diffuse gas and dust structures in the Galaxy, so giving them separate degrees of freedom is expected to have a negligible effect on the results.

We highlight the fact that the IC and π^0+Bremsstrahlung templates are allowed to vary separately in the Model A, B, and C fits. As a result, the foreground templates in these tests are given considerable freedom in the fitting procedure, as they are associated with sixteen free parameters (rather than just eight, as in the `p6v11` case). This is a very important cross-check of the fiducial results, because the relative normalizations of foreground components are fixed in `p6v11`, with the ratio set by a previous fit to the data. However, because that fit did not include a DM template, one might worry that a potential signal—if present—would be absorbed by the foreground components (particularly the IC component) in the initial fitting procedure. If this were the case, using `p6v11` for a Galactic DM search could potentially give artificially stringent DM limits.

Fig. 3 shows the limits obtained using Models A, B, and C. The differences between the results can be understood in terms of the assumptions going into the separate models, which we now describe in detail:

Model A is based on the class of Galactic diffuse models studied in [68], and is described in detail in [8]. Here, we only highlight the main elements that distinguish it

from Models B and C. For Model A, cosmic-ray electrons and nuclei are both sourced by the same population of pulsars, and the cosmic-ray diffusion coefficient and re-acceleration strength are held constant. The left panel of Fig. 3 shows the Monte Carlo expectation and data limit when rerunning the fiducial analysis using the Model A templates. The recovered data limit is weaker than the Monte Carlo expectation, which suggests that there is excess gamma-ray emission in the ROI that is not captured by the Model A templates. It should be noted that the foreground templates are given considerable freedom in the fitting procedure, as the normalizations of the π^0+Bremsstrahlung and IC templates are allowed to float separately in each radial slice. Despite this freedom, a large amount of DM emission is still needed to improve the quality of the fit. A DM "excess" with a $\mathrm{TS_{max}} \sim 28$ is observed, with the best-fit 1σ and 2σ (corresponding to deviations in TS of -2.30 and -6.18 from the global maximum) containment regions as shown in the figure. The fact that the DM parameter space that is favored is clearly excluded by the dwarf searches strongly suggests that the weakening of the bounds is not due to DM, and is likely of astrophysical origin.

Model B provides an important counterpoint to Model A [8]. It includes an additional source population of electrons at the Galactic Center, which contributes to the IC emission. Unlike Model A, which closely reproduces the local cosmic-ray electron spectrum, Model B under-predicts the distribution below ~ 20 GV. However, this disparity can be accounted for by contributions from other more local sources. The middle panel of Fig. 3 shows the Monte Carlo expectation and data limit for the Model B study. The limit is comparable to the fiducial case at low masses and is somewhat tighter for masses above ~ 100 GeV, although still consistent within the
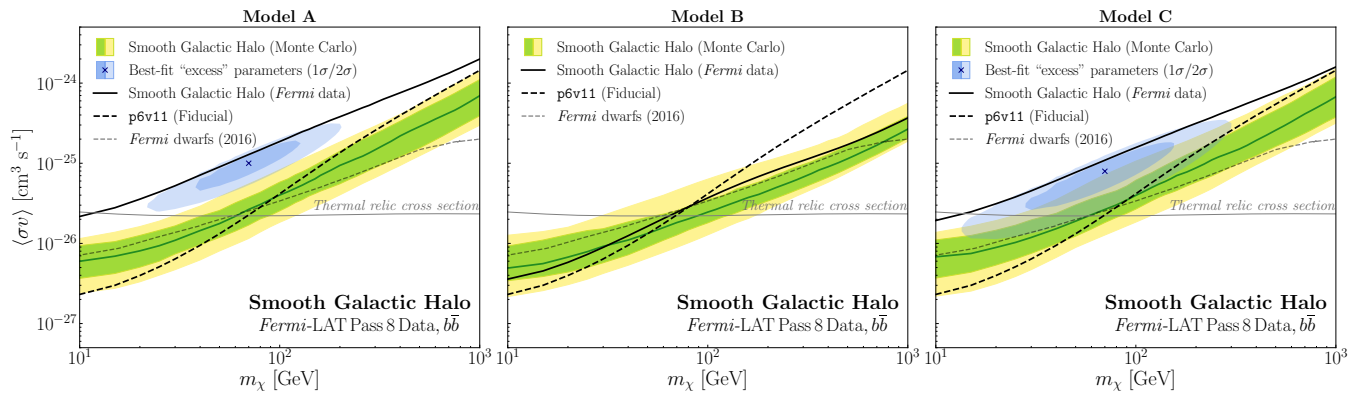
FIG. 3. Similar to Fig. 1 of the paper, except using the Model A, B, and C foreground models (left, middle, and right panel, respectively) as provided by the *Fermi*-LAT Collaboration [8]. Note that the foreground templates are still divided into eight radial slices, as in the fiducial study, but the normalizations of the inverse-Compton and π^0+Bremsstrahlung templates are allowed to float independently. The fiducial limit obtained using the `p6v11` foreground model is shown by the dashed black line for comparison. The "excesses" in the Model A and C studies (with significances $TS_{max} \sim 28$ and 14, respectively) are well-understood in terms of the source populations included in these models; see text for further discussion. Model B is statistically preferred over Models A and C as a description of the data in our ROI; the difference in the maximum log-likelihood between Model B and A(C) is $\Delta \log \mathcal{L}_{max} = 136(119)$.

Monte Carlo expectation. The predicted IC spectrum from `Galprop` that is used in Model B tends to be a better match to the fitted spectrum (compared to Models A and C). The better overall fit of Model B to the data in this case and the fact that the additional emission is absorbed by the IC template means that an astrophysical origin of the excess is statistically preferred to the DM component.

For Model C, the cosmic-ray diffusion coefficient and re-acceleration strength depends on the Galactocentric radius and height [8]. Additionally, while the cosmic-ray electron/nuclei are sourced from the same population, their distribution is more central than that used for Model A. The differences between Model A and C predominantly show up in the outer galaxy, and so the two give largely similar results when used within our ROI. The excess emission observed in the case of Model A is also present using Model C, with a preference for roughly similar DM parameter values. Again, the fact that the preferred parameter space is robustly ruled out by dwarf searches strongly indicates that the excess emission in this case is of astrophysical origin.

To summarize, Model B provides limits very similar to those obtained in the fiducial case, while Models A and C exhibit significant excesses above Monte Carlo expectation. This difference can be attributed to the fact that Model B includes an additional population of electron-only sources near the Galactic Center that contributes to the IC emission. Omission of this population in Models A and C causes the DM template to absorb more flux, thus weakening the overall bounds. Overall, the fitted IC normalization for Model B is closer to its initial `Galprop` prediction (with a value ~ 1.1), as compared to that for Models A and C (with a value ~ 2.4) [8]. This suggests that, of the three scenarios considered, Model B may best

capture the IC emission in the ROI used here.

C. The GeV Excess

The results presented in this paper have direct implications for the interpretation of the excess of GeV photons observed in the Galactic Center. If the GeV excess arises from DM, then the signal should also contribute a photon flux in the ROI studied here. This is a more direct comparison than using dwarf galaxies or galaxy groups because it removes uncertainties having to do with differences in halo density distribution. In Fig. 4, we show the Galactic DM limits obtained for various assumptions of the inner slope, γ, of the generalized NFW density profile. The steeper the inner slope, the stronger the annihilation limit. Results are shown for annihilation into $b\bar{b}$ (left) and $\tau^+\tau^-$ (right). For comparison, we also show the best-fit regions to the Galactic Center gamma-ray excess from previous work as the data point [18] and solid [15], cross-hatched [19], and hatched [21] regions. The DM interpretation of the GeV excess typically prefers a steeper inner slope with $\gamma \gtrsim 1.1$, where the limits from the Galactic halo become quite stringent. These Milky Way limits robustly exclude a DM interpretation of the excess for the $b\bar{b}$ channel, and, for the first time, start probing the $\tau^+\tau^-$ scenario. Explanations in terms of other annihilation channels are also highly constrained, as reviewed in App. C.

Variations in the foreground modeling can affect the recovered limits from our analysis and their implications for the Galactic Center Excess. Of the variations explored in Sec III B, Model B appears to best capture the IC emission in the ROI, as the fitted normalization of
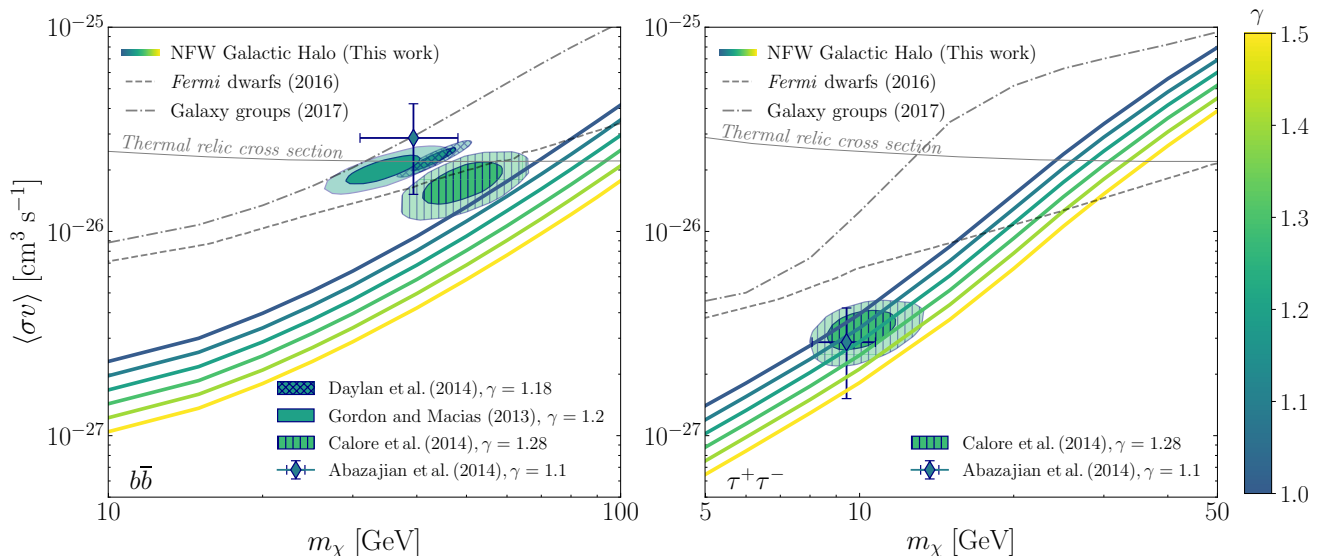
FIG. 4. The 95% confidence limits on dark matter annihilation into $b\bar{b}$ (left) and $\tau^+\tau^-$ (right) for the fiducial analysis, varying over the inner slope, γ, of the generalized NFW density profile. The limits tighten as γ increases; the lines shown correspond to linearly spaced steps from $\gamma = 1$ to 1.5. The best-fit parameters obtained by previous studies of the GeV excess are indicated by the data point [18] and the solid [15], cross-hatched [19], and hatched [21] regions. Each region is indicated by $1\sigma/2\sigma$ contours and colored corresponding to the best-fit γ obtained by that study, also specified in the legend. For ease of comparison, we have rescaled the best-fit cross-sections to be consistent with $\rho(r_\odot) = 0.4 \, \text{GeV cm}^{-3}$. The corresponding limits obtained from dwarf galaxies [5] and galaxy groups [7] (grey dashed and dot-dashed lines, respectively) are also shown. The expected annihilation cross section for a generic weakly interacting massive particle is indicated by the solid grey line [67].

this component is closest to its initial `Galprop` value. The limits obtained using Model B are only marginally weaker than those using `p6v11` at low masses and still robustly disfavor the DM interpretation of the excess in terms of annihilation into the $b\bar{b}$ final state. These results are suggestive, but do not eliminate the systematic uncertainties associated with diffuse emission modeling. To sidestep this issue, we can choose to compare our results to only those Inner Galaxy studies that use the same Galactic foreground model as we do. The cross-hatched region in Fig. 4 is derived using the `p6v11` diffuse foreground model [19] and therefore provides the most direct comparison to our limit. It is strongly excluded by the limit we recover for the corresponding value of γ.

variations in terms of the different physical assumptions underlying the foreground models. We have performed rigorous Monte Carlo and injected signal tests to ensure the robustness of our results. This study excludes the $b\bar{b}$ annihilation interpretation of the Galactic Center excess at 95% confidence for the `p6v11` diffuse model, and for the first time starts probing the $\tau^+\tau^-$ annihilation interpretation.

The Appendices complement the discussion here with extended results. In particular, App. A includes further justification for the choice of ROI and App. B summarizes signal injection and recovery tests. Extended results, including limits for different annihilation channels and DM profiles, as well as other variations of the astrophysical templates, are provided in App. C.

IV. CONCLUSIONS

In this paper, we have presented a comprehensive search for DM annihilation from the smooth Milky Way halo in *Fermi* gamma-ray data. We do not find significant evidence for an annihilation signal, and obtain strong bounds on the properties of annihilating DM. We exclude thermal dark matter at masses below $\sim$70 GeV for the $b\bar{b}$ annihilation channel when using the *Fermi* `p6v11` diffuse model, representing the strongest limits to date in this mass range. We have carefully considered uncertainties associated with the modeling of the diffuse Galactic foregrounds and are able to understand these

V. ACKNOLWEDGEMENTS

[1] W. B. Atwood et al. (Fermi-LAT), Astrophys. J. **697**, 1071 (2009), 0902.1089.

[2] M. Ackermann et al. (Fermi-LAT), Phys. Rev. **D89**, 042001 (2014), 1310.0828.

[3] A. Geringer-Sameth, S. M. Koushiappas, and M. G. Walker, Phys. Rev. **D91**, 083535 (2015), 1410.2242.

[4] M. Ackermann et al. (Fermi-LAT), Phys. Rev. Lett. **115**, 231301 (2015), 1503.02641.

[5] A. Albert et al. (DES, Fermi-LAT), Astrophys. J. **834**, 110 (2017), 1611.03184.

[6] M. Lisanti, S. Mishra-Sharma, N. L. Rodd, B. R. Safdi, and R. H. Wechsler (2017), 1709.00416.

[7] M. Lisanti, S. Mishra-Sharma, N. L. Rodd, and B. R. Safdi (2017), 1708.09385.

[8] M. Ackermann et al. (Fermi-LAT), Astrophys. J. **799**, 86 (2015), 1410.3696.

[9] L. Goodenough and D. Hooper (2009), 0910.2998.

[10] D. Hooper and L. Goodenough, Phys. Lett. **B697**, 412 (2011), 1010.2752.

[11] A. Boyarsky, D. Malyshev, and O. Ruchayskiy, Phys. Lett. **B705**, 165 (2011), 1012.5839.

[12] D. Hooper and T. Linden, Phys. Rev. **D84**, 123005 (2011), 1110.0006.

[13] K. N. Abazajian and M. Kaplinghat, Phys. Rev. **D86**, 083511 (2012), [Erratum: Phys. Rev.D87,129902(2013)], 1207.6047.

[14] D. Hooper and T. R. Slatyer, Phys. Dark Univ. **2**, 118 (2013), 1302.6589.

[15] C. Gordon and O. Macias, Phys. Rev. **D88**, 083521 (2013), [Erratum: Phys. Rev.D89,no.4,049901(2014)], 1306.5725.

[16] W.-C. Huang, A. Urbano, and W. Xue (2013), 1307.6862.

[17] O. Macias and C. Gordon, Phys. Rev. **D89**, 063515 (2014), 1312.6671.

[18] K. N. Abazajian, N. Canac, S. Horiuchi, and M. Kaplinghat, Phys. Rev. **D90**, 023526 (2014), 1402.4090.

[19] T. Daylan, D. P. Finkbeiner, D. Hooper, T. Linden, S. K. N. Portillo, N. L. Rodd, and T. R. Slatyer, Phys. Dark Univ. **12**, 1 (2016), 1402.6703.

[20] B. Zhou, Y.-F. Liang, X. Huang, X. Li, Y.-Z. Fan, L. Feng, and J. Chang, Phys. Rev. **D91**, 123010 (2015), 1406.6948.

[21] F. Calore, I. Cholis, and C. Weniger, JCAP **1503**, 038 (2015), 1409.0042.

[22] K. N. Abazajian, N. Canac, S. Horiuchi, M. Kaplinghat, and A. Kwa, JCAP **1507**, 013 (2015), 1410.6168.

[23] M. Ajello et al. (Fermi-LAT), Astrophys. J. **819**, 44 (2016), 1511.02938.

[24] C. Karwin, S. Murgia, T. M. P. Tait, T. A. Porter, and P. Tanedo, Phys. Rev. **D95**, 103005 (2017), 1612.05687.

[25] S. K. Lee, M. Lisanti, B. R. Safdi, T. R. Slatyer, and W. Xue, Phys. Rev. Lett. **116**, 051103 (2016), 1506.05124.

[26] R. Bartels, S. Krishnamurthy, and C. Weniger, Phys. Rev. Lett. **116**, 051102 (2016), 1506.05104.

[27] O. Macias, C. Gordon, R. M. Crocker, B. Coleman, D. Paterson, S. Horiuchi, and M. Pohl (2016), 1611.06644.

[28] R. Bartels, E. Storm, C. Weniger, and F. Calore (2017), 1711.04778.

[29] R. Keeley, K. Abazajian, A. Kwa, N. Rodd, and B. Safdi (2017), 1710.03215.

[30] J. L. Sanders, N. W. Evans, A. Geringer-Sameth, and W. Dehnen, Phys. Rev. **D94**, 063521 (2016), 1604.05493.

[31] V. Bonnivard, D. Maurin, and M. G. Walker, MNRAS **462**, 223 (2016), 1506.08209.

[32] A. Geringer-Sameth, S. M. Koushiappas, and M. Walker, Astrophys. J. **801**, 74 (2015), 1408.0002.

[33] K. N. Abazajian and R. E. Keeley, Phys. Rev. **D93**, 083514 (2016), 1510.06424.

[34] M. Papucci and A. Strumia, JCAP **1003**, 014 (2010), 0912.0742.

[35] M. Cirelli, P. Panci, and P. D. Serpico, Nucl. Phys. **B840**, 284 (2010), 0912.0663.

[36] E. J. Baxter and S. Dodelson, Phys. Rev. **D83**, 123516 (2011), 1103.5779.

[37] D. Malyshev, J. Bovy, and I. Cholis, Phys. Rev. **D84**, 023013 (2011), 1007.4556.

[38] M. Ackermann et al. (Fermi-LAT), Astrophys. J. **761**, 91 (2012), 1205.6474.

[39] H. S. Zechlin, S. Manconi, and F. Donato (2017), 1710.01506.

[40] X. Huang, T. Enßlin, and M. Selig, JCAP **1604**, 030 (2016), 1511.02621.

[41] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, Astrophys. J. **622**, 759 (2005), astro-ph/0409513.

[42] D. Gaggero, M. Taoso, A. Urbano, M. Valli, and P. Ullio, JCAP **1512**, 056 (2015), 1507.06129.

[43] M. Werner, R. Kissmann, A. W. Strong, and O. Reimer, Astropart. Phys. **64**, 18 (2015), 1410.5266.

[44] R. Kissmann, Astropart. Phys. **55**, 37 (2014), 1401.4035.

[45] D. Gaggero, A. Urbano, M. Valli, and P. Ullio, Phys. Rev. **D91**, 083012 (2015), 1411.7623.

[46] C. Evoli, D. Gaggero, A. Vittino, G. Di Bernardo, M. Di Mauro, A. Ligorini, P. Ullio, and D. Grasso, JCAP **1702**, 015 (2017), 1607.07886.

[47] E. Storm, C. Weniger, and F. Calore, JCAP **1708**, 022 (2017), 1705.04065.

[48] M. Selig, V. Vacca, N. Oppermann, and T. A. Enßlin, Astron. Astrophys. **581**, A126 (2015), 1410.4562.

[49] M. Su, T. R. Slatyer, and D. P. Finkbeiner, Astrophys. J. **724**, 1044 (2010), 1005.5480.

[50] F. Acero et al. (Fermi-LAT), Astrophys. J. Suppl. **218**, 23 (2015), 1501.02003.

[51] J. F. Navarro, C. S. Frenk, and S. D. M. White, Astrophys. J. **490**, 493 (1997), astro-ph/9611107.

[52] C. F. McKee, A. Parravano, and D. J. Hollenbach, Astrophys. J. **814**, 13 (2015), 1509.05334.

[53] S. Sivertsson, H. Silverwood, J. I. Read, G. Bertone, and P. Steger, Submitted to: Mon. Not. Roy. Astron. Soc. (2017), 1708.07836.

[54] J. I. Read, J. Phys. **G41**, 063101 (2014), 1404.1938.

[55] A. W. Strong, I. V. Moskalenko, and O. Reimer, Astrophys. J. **537**, 763 (2000), [Erratum: Astrophys. J.541,1109(2000)], astro-ph/9811296.

[56] F. Acero et al. (Fermi-LAT), Astrophys. J. Suppl. **223**, 26 (2016), 1602.07246.

[57] M. I. Large, M. J. S. Quigley, and C. G. T. Haslam, MNRAS **124**, 405 (1962).

[58] C. G. T. Haslam, U. Klein, C. J. Salter, H. Stoffel, W. E. Wilson, M. N. Cleary, D. J. Cooke, and P. Thomasson, Astron. Astrophys. **100**, 209 (1981).

[59] J.-M. Casandjian, I. Grenier, and for the Fermi Large Area Telescope Collaboration, ArXiv e-prints (2009), 0912.3478.

[60] M. Ackermann et al. (Fermi-LAT), Phys. Rev. **D85**, 083007 (2012), 1202.2856.

[61] M. Ackermann et al. (Fermi-LAT), Astrophys. J. **793**, 64 (2014), 1407.7905.

[62] W. A. Rolke, A. M. Lopez, and J. Conrad, Nucl. Instrum. Meth. **A551**, 493 (2005), physics/0403059.

[63] M. Cirelli, G. Corcella, A. Hektor, G. Hutsi, M. Kadastik, P. Panci, M. Raidal, F. Sala, and A. Strumia, JCAP **1103**, 051 (2011), [Erratum: JCAP1210,E01(2012)], 1012.4515.

[64] S. Mishra-Sharma, N. L. Rodd, and B. R. Safdi, Astron. J. **153**, 253 (2017), 1612.03173.

[65] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, SIAM J. Sci. Comput. **16**, 1190 (1995), ISSN 1064-8275, URL http://dx.doi.org/10.1137/0916069.

[66] E. Jones, T. Oliphant, P. Peterson, et al., *SciPy: Open source scientific tools for Python* (2001–), URL http://www.scipy.org/.

[67] G. Steigman, B. Dasgupta, and J. F. Beacom, Phys. Rev. **D86**, 023506 (2012), 1204.3622.

[68] M. Ackermann, M. Ajello, W. B. Atwood, L. Baldini, J. Ballet, G. Barbiellini, D. Bastieri, K. Bechtol, R. Bellazzini, B. Berenji, et al., Astrophys. J. **750**, 3 (2012), 1202.4039.

[69] T. Cohen, K. Murase, N. L. Rodd, B. R. Safdi, and Y. Soreq, Phys. Rev. Lett. **119**, 021102 (2017), 1612.05638.

[70] Astropy Collaboration, T. P. Robitaille, E. J. Tollerud, P. Greenfield, M. Droettboom, E. Bray, T. Aldcroft, M. Davis, A. Ginsburg, A. M. Price-Whelan, et al., AAP **558**, A33 (2013), 1307.6212.

[71] F. Pérez and B. E. Granger, Computing in Science and Engineering **9**, 21 (2007), ISSN 1521-9615, URL http://ipython.org.

[72] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[73] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. **C71**, 1554 (2011), [Erratum: Eur. Phys. J.C73,2501(2013)], 1007.1727.

[74] J. Einasto, Trudy Astrofizicheskogo Instituta Alma-Ata **5**, 87 (1965).

[75] J. F. Navarro, A. Ludlow, V. Springel, J. Wang, M. Vogelsberger, S. D. M. White, A. Jenkins, C. S. Frenk, and A. Helmi, Mon. Not. Roy. Astron. Soc. **402**, 21 (2010), 0810.1522.

[76] A. Burkert, IAU Symp. **171**, 175 (1996), [Astrophys. J.447,L25(1995)], astro-ph/9504041.

[77] A. V. Maccio', G. Stinson, C. B. Brook, J. Wadsley, H. M. P. Couchman, S. Shen, B. K. Gibson, and T. Quinn, Astrophys. J. **744**, L9 (2012), 1111.5620.

[78] A. Di Cintio, C. B. Brook, A. V. Macciò, G. S. Stinson, A. Knebe, A. A. Dutton, and J. Wadsley, Mon. Not. Roy. Astron. Soc. **437**, 415 (2014), 1306.0898.

[79] F. Calore, I. Cholis, C. McCabe, and C. Weniger, Phys. Rev. **D91**, 063003 (2015), 1411.4647.

Appendix A: The Region-of-Interest

The fiducial analysis presented in this paper uses a region-of-interest (ROI) defined by the annulus $|b| > 20°$ and $r < 50°$. To motivate this choice, we analyze Asimov datasets [73], which can be used to determine the median asymptotic behavior of the test statistic under the assumption that the foregrounds are perfectly modeled, while varying over different choices of the ROI. The Asimov dataset in this case corresponds to the sum of astrophysical templates best-fit to the data in each ROI. Note that the p6v11 template was not divided into independent radial slices in the Asimov study.

As a concrete example, we consider the case of a 30 GeV DM particle annihilating to $b\bar{b}$, although results for other DM masses are largely unchanged. We vary over latitude ($|b| > b_{\mathrm{cut}}$) and radial ($r < r_{\mathrm{cut}}$) cuts spanning $b_{\mathrm{cut}} = \{15°, 16°, \ldots, 30°\}$ and $r_{\mathrm{cut}} = \{40°, 45°, \ldots, 150°\}$. Figure 5 demonstrates how the projected cross section limit, $\langle \sigma v \rangle_{\mathrm{lim}}$, compares to that for the fiducial ROI, $\langle \sigma v \rangle_{\mathrm{lim}}^{\mathrm{fid}}$, as a function of b_{cut} and r_{cut}. We consider the generalized NFW profile as in Eq. 3 with scale radius $r_s = 17$ kpc, local density $\rho(r_\odot) = 0.4$ GeV cm^{-3}, and inner slope $\gamma = 1$ and 1.2 (first and second panel from left, respectively). In general, we see that the projected sensitivity strengthens for smaller b_{cut} and larger r_{cut}, as expected. This dependence weakens for steeper profiles because the dark matter (DM) density is concentrated towards the Galactic Center. We note that the Asimov projections assume perfect knowledge of the astrophysical components, and as such disregard potential degeneracies between a DM signal and astrophysical templates, which are likely to be important in an analysis on data.

For comparison, we also consider several other DM density profiles, each normalized to $\rho(r_\odot) = 0.4$ GeV cm^{-3}. The middle panel of Fig. 5 shows the results for the Einasto profile [74]:

$$\rho_{\mathrm{Einasto}}(r) = \rho_0 \exp\left[-\frac{2}{\alpha} \left(\left(\frac{r}{r_E} \right)^\alpha - 1 \right) \right], \quad \text{(A1)}$$

with $\alpha = 0.17$ and $r_E = 15.14$ kpc [75]. The final two panels in Fig. 5 show the results for a cored Burkert profile [76]:

$$\rho_{\mathrm{Burkert}}(r) = \frac{\rho_0}{(1 + r/r_B)[1 + (r/r_B)^2]}, \quad \text{(A2)}$$

where r_B is the analog the NFW scale radius and sets the size of the core. For illustration, we consider $r_B = 0.5$ and 10 kpc, which roughly spans the range of allowed
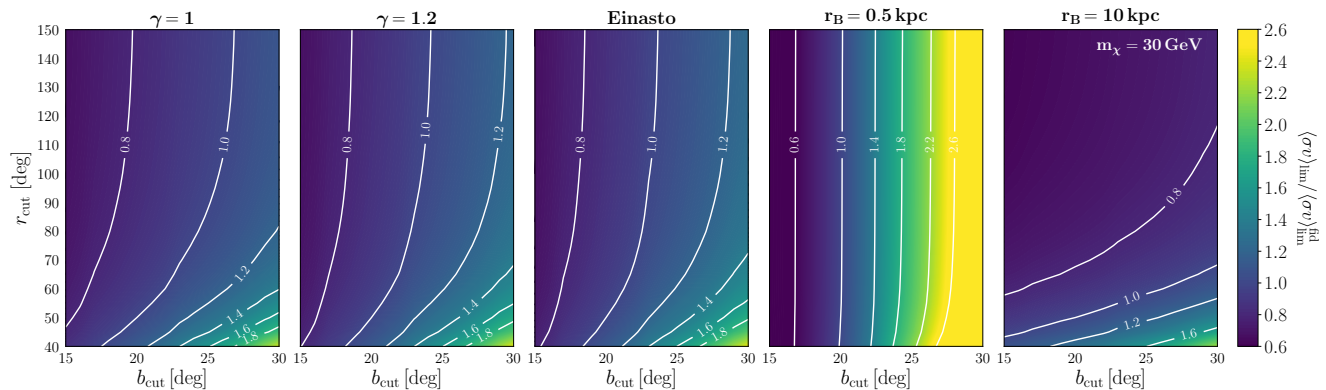
FIG. 5. Sensitivity projections for a 30 GeV dark matter particle annihilating to $b\bar{b}$ for different regions of interest, which are defined by latitude ($|b| > b_{\rm cut}$) and radial ($r < r_{\rm cut}$) cuts. The projected limit, $\langle \sigma v \rangle_{\rm lim}$, is compared to the limit for the fiducial region, $\langle \sigma v \rangle_{\rm lim}^{\rm fid}$, which corresponds to $|b| > 20°$ and $r < 50°$. The contours indicate the ratio of these two cross sections. The projections are provided for different dark matter density profiles: (left to right) generalized NFW with inner-slope $\gamma = 1, 1.2$, Einasto, and Burkert with a $r_B = 0.5$ and 10 kpc core.

possibilities—see *e.g.* [77, 78]. While the Einasto contours look very similar to those for NFW with $\gamma = 1.2$, the Burkert results are quite different. For the smaller core, there is only very mild dependence on $r_{\rm cut}$ and the projected signal strength decreases with larger $b_{\rm cut}$. In contrast, the signal is strengthened with decreased latitude and increased radial cuts for the case where $r_B = 10$ kpc because the DM distribution is less concentrated towards the Galactic Center.

Appendix B: Signal Injection and Recovery

A vital consistency check involves ensuring that the limit-setting procedure would not exclude a DM signal if one were present in the data. We perform a variety of tests to confirm that we can set a robust limit while recovering the properties of a DM signal. We perform these checks on both Monte Carlo simulations as well as on the data itself.

Signal injection on Monte Carlo. We create Monte Carlo simulations of the gamma-ray sky by summing the astrophysical templates best-fit on data, adding the signal from a DM particle annihilating to $b\bar{b}$ in the smooth Galactic halo, and Poisson fluctuating the final map. We create 50 Monte Carlo realizations of the sky map and pass each through the analysis pipeline. This procedure is repeated for different DM masses and cross sections to study the resulting limit and the test statistic associated with the extracted signal.

Figure 6 summarizes the results of the signal injection tests for $m_\chi = 100$ and 1000 GeV in the left and right panel, respectively. In each panel, the gold bands indicate the recovered limit, $\langle \sigma v \rangle_{\rm limit}^{\rm null}$, when no signal is injected into the simulated sky map. The green band

shows the middle 68% containment of the cross section, $\langle \sigma v \rangle_{\rm limit}^{\rm inj}$, that is recovered when TS $= -2.71$ in the presence of an injected signal with cross section $\langle \sigma v \rangle_{\rm inj}$. If the statistical procedure is robust, the green band should lie above the diagonal line (saying that the limit set would be consistent with an injected signal) and should asymptotically approach the gold band for small signal cross sections, as is indeed the case for both masses included here.

The blue line shows the recovered cross section that is associated with the maximum test statistic, $\rm TS_{max}$:

$$\mathrm{TS}_{\max} \equiv 2\left[\log \mathcal{L}(d|\mathcal{M}, \widehat{\langle \sigma v \rangle}, m_\chi)\right.$$
$$\left. - \log \mathcal{L}(d|\mathcal{M}, \langle \sigma v \rangle = 0, m_\chi)\right], \quad \text{(B1)}$$

where $\widehat{\langle \sigma v \rangle}$ is the cross section that maximizes the likelihood. In the regime where $\rm TS_{max} < 1$, this is shown as a dashed line. The blue band corresponds to the range of cross sections above and below $\widehat{\langle \sigma v \rangle}$ associated with $\rm TS_{max} - 1$, spanning the extremal values of the middle 68% containment in each case. We expect that the recovered cross section should be consistent with statistical noise once the limit is reached, as is clearly demonstrated. The inset in each panel of Fig. 6 demonstrates how $\rm TS_{max}$ depends on the injected cross section.

While we show the representative cases for DM masses $m_\chi = 100$ and 1000 GeV here, we find that signal injection tests on Monte Carlo are well-behaved for DM masses ranging from 10–1000 GeV. The tests fail when the upper cutoff on the photon energy is $\gtrsim 100$ GeV most likely due to limited photon statistics. For this reason, as well as the fact that the p6v11 template should be used with caution at energies $\gtrsim 50$ GeV, we have restricted the photon energies to be below $\sim$50 GeV.

Signal injection on data. We also perform a data-driven version of the signal injection tests, adding a
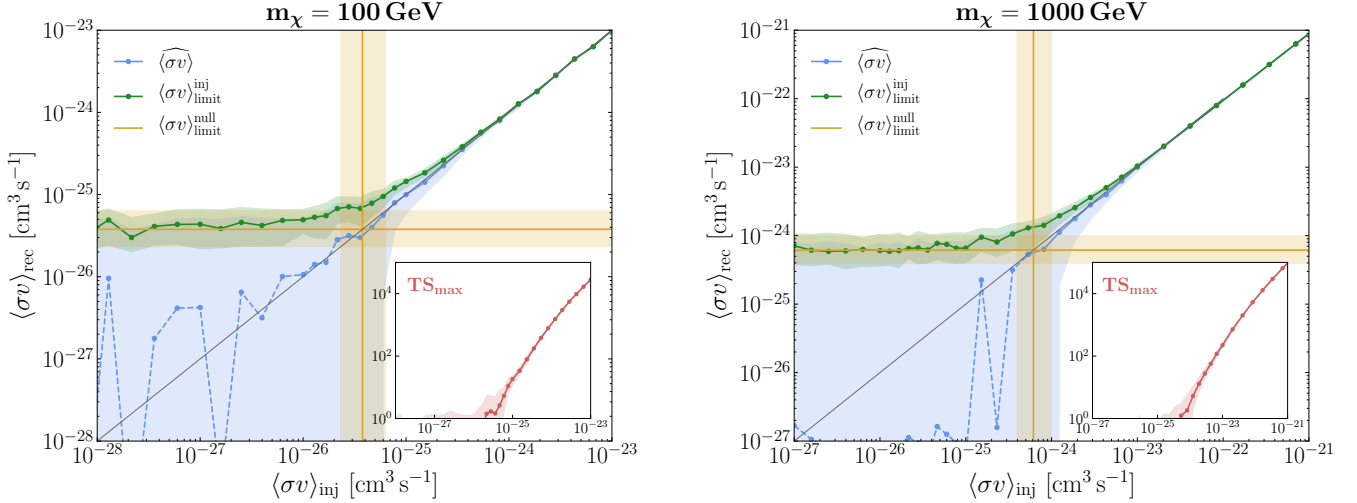
FIG. 6. Signal injection tests on Monte Carlo simulations for a 100 (left) and 1000 (right) GeV DM particle annihilating to $b\bar{b}$. In each panel, the gold line corresponds to the limit $\langle\sigma v\rangle_{\text{limit}}^{\text{null}}$ obtained when no signal is injected into the simulated data. The green line corresponds to the median cross section limit, $\langle\sigma v\rangle_{\text{limit}}^{\text{inj}}$, that is recovered for a given injected cross section $\langle\sigma v\rangle_{\text{inj}}$, when TS $= -2.71$. The green band shows the corresponding 68% containment. The blue line corresponds to the median recovered cross section $\widehat{\langle\sigma v\rangle}$ that is associated with the maximum test statistic TS_{max} (plotted in the inset), and is shown as dashed in the regime where $\text{TS}_{\text{max}} < 1$. The blue band spans extremal values of the 68% containment of cross sections associated with $\text{TS}_{\text{max}} - 1$. For each injected signal point, we create 50 realizations of simulated sky maps.
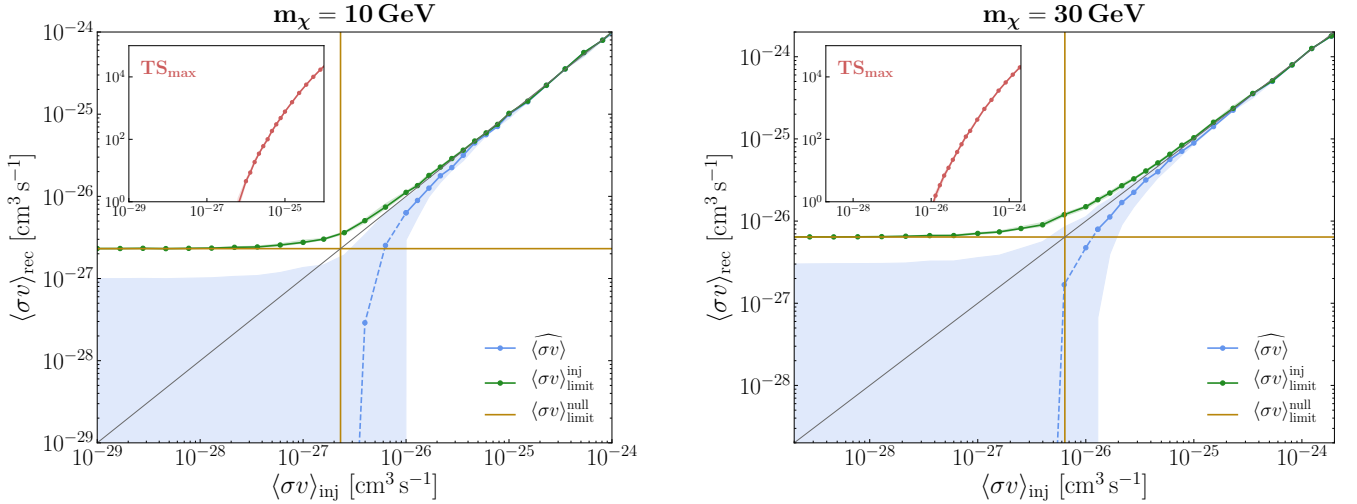

FIG. 7. The same as Fig. 6, except for signal injected on data. The left(right) panel corresponds to a 10(30) GeV DM mass. In this case, for each injected signal point, we create 10 realizations of simulated sky maps.

Galactic DM signal for $b\bar{b}$ annihilation on top of the actual data and passing this through the analysis pipeline. We repeat this procedure for 10 sky map realizations. This is a particularly important check at lower energies, where effects of point spread function (PSF) and foreground mis-modeling can lead to artificially strong limits for lower DM masses. Figure 7 summarizes the results of the signal injection tests on data for DM masses of 10 and 30 GeV (left and right panel, respectively). In

each case, we see that the analysis would not exclude an injected DM signal. We restrict ourselves to energies $E_\gamma \gtrsim 0.8$ GeV to mitigate the effects of a significantly degraded PSF at even lower energies. We caution that while this procedure demonstrates that a signal would not be excluded under the null assumption on the data, it is still possible that mis-modeling effects can impact the final result for the lowest masses (~ 10 GeV). This can be seen in the left panel of Fig. 7 from the fact that

the median recovered cross section associated with $\mathrm{TS}_{\max}$ (blue line) falls and becomes consistent with zero slightly above the null limit. However, this small discrepancy occurs in the range where $\mathrm{TS}_{\max} \lesssim 1$. This is not an issue for higher masses; for example, for a DM mass of 30 GeV, the median recovered cross section associated with $\mathrm{TS}_{\max}$ is consistent with zero only for cross sections below the null limit, as shown in the right panel of Fig. 7.

We have also performed signal injection tests using Model B to ensure the validity of the recovered bounds. The left panel of Fig. 8 shows the results of signal injection on data, as described in the previous section, for a DM mass of $m_\chi = 10$ GeV where foreground and PSF mis-modeling are likely to have the largest effect. We see that a putative DM signal would not be excluded by the analysis in this case. We also show results for an injected DM mass of $m_\chi = 70$ GeV in the right panel of Fig. 8, corresponding to the value most consistent with the excess emission seen in Models A and C. Again, we see that a potential DM signal would not be excluded in this case.

Appendix C: Extended Results

We consider several additional variations to the fiducial analysis, and summarize the results here:

- Although we presented results for DM annihilating into the $b\bar{b}$ and $\tau^+\tau^-$ final states in the paper, DM annihilation can proceed into a variety of Standard Model final states. In Fig. 9 (left), we reinterpret the main results of the fiducial study in terms of annihilation into additional final states. Broadly, the spectra of hadronic channels (W^+W^-, ZZ, $q\bar{q}$, $c\bar{c}$, $t\bar{t}$) are predominantly set by boosted π^0 decays, resulting in comparable final limits beyond the respective mass thresholds. Gamma-rays for the leptonic ($e^+e^-, \mu^+\mu^-$) channels predominantly arise from radiative decays and final-state radiation, resulting in somewhat weaker overall limits. In each case, we assume 100% branching fraction into the specified channel. Note that we only model prompt gamma-ray emission and do not account for inverse-Compton or synchrotron radiation of the final state [63], which is relevant for the lighter leptonic channels.

- In addition to the $b\bar{b}$ and $\tau^+\tau^-$ cases considered in the paper, we summarize in Fig. 9 (right) constraints on other possible annihilation channels contributing to the GeV excess. We show our results for the $q\bar{q}, c\bar{c}, gg$ and hh final states, spanning the range $\gamma = 1.2$–1.3 for the inner slope of the NFW generalized profile (thick bands), along with the corresponding best-fit contours as found by [79] assuming $\gamma = 1.28$. We see that the $q\bar{q}$ and hh explanations are robustly excluded by this analysis,

while the $c\bar{c}$ and gg explanations are put significantly under tension. We do not include annihilation channels that are already excluded at the 95% confidence level by spectral fits to the *Fermi* GeV excess emission [79].

- Figure 4 of the paper demonstrates how the fiducial limit depends on the inner slope of the NFW profile. We have additionally considered the Einasto and Burkert profiles, defined in Eq. A1 and A2. The associated limits are shown in Fig. 10. The Einasto limit (solid green) is a factor of $\lesssim 1.6$ stronger than the fiducial case, while the Burkert limit is a factor of $\lesssim 24(5)$ stronger(weaker) for $r_B = 0.5(10)$ kpc (dotted and dashed green, respectively).

- We assumed a local DM density of $\rho(r_\odot) = 0.4$ GeV cm^{-3} in the fiducial analysis, consistent with recent measurements [52, 53]. Other estimates in the literature, however, point to a value closer to $\rho(r_\odot) = 0.3$ GeV cm^{-3} (see [54] and references therein). Repeating the analysis using this lower value, we find that the limit is $\lesssim 1.8$ times weaker (solid blue line in Fig. 10). We emphasize that the assumption made about the local DM density does not impact the conclusions drawn about the viability of the GeV excess, as the best-fit regions are similarly shifted to higher annihilation cross sections by roughly the same amount.

- Our fiducial analysis does not account for potential emission from Loop I in the Northern hemisphere. As a proxy for this contribution, we include an additional isotropic template in the Northern hemisphere. Modeling this emission results in a slight improvement in the DM constraint by a factor of $\lesssim 1.2$ (dashed purple line in Fig. 10), as expected because additional foreground components are accounted for.

- In the fiducial study, the Northern and Southern lobes of the *Fermi* bubbles are floated separately. We have verified that floating the Northern and Southern lobes together leave the limit unchanged. Figure 10 shows what happens if the *Fermi* bubbles are not included at all in the analysis. In this case, the limit worsens by a factor of $\lesssim 6$ (solid gold line).

- In the fiducial study, all point sources were masked to 95% containment in PSF, according to energy bin. To estimate the effect of point-source mis-modeling, we increased the mask size to 99% PSF containment; this results in a factor of $\lesssim 1.5$ weakening of the fiducial limit (solid purple line in Fig. 10), likely due to the corresponding reduction in the effective size of the ROI.

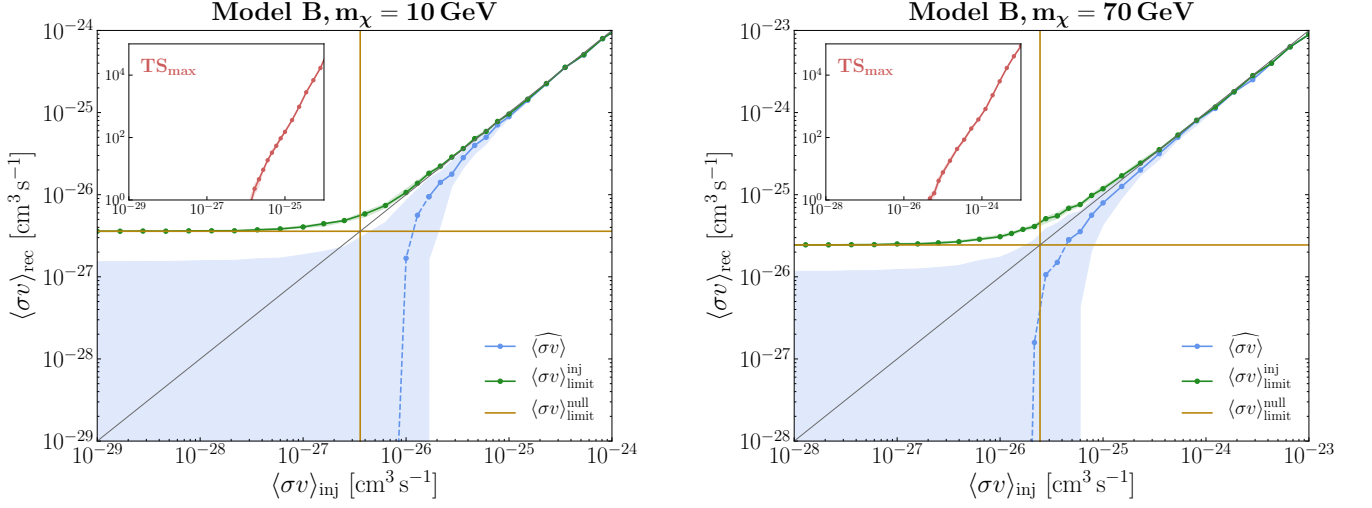- The fiducial analysis takes full advantage of the spatial profiles of the expected DM emission and

FIG. 8. Signal injection test on data using the Model B foreground template, assuming $m_\chi = 10$ GeV (left) and 70 GeV (right). Format as in Fig. 7.
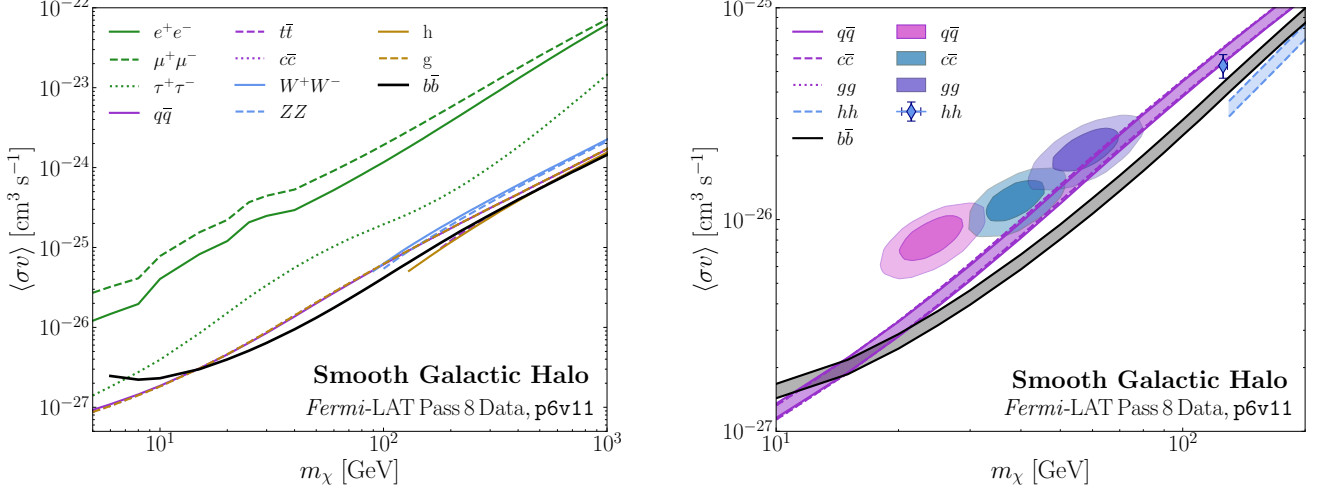

FIG. 9. (Left) The 95% confidence limit on dark matter of mass, m_χ, annihilating with cross section, $\langle \sigma v \rangle$, in the smooth Galactic halo. The limits are obtained following the fiducial analysis procedure described in the paper, but varying over the annihilation channel. (Right) The 95% confidence limits on dark matter annihilation into $b\bar{b}$ (fiducial), $q\bar{q}$, $c\bar{c}$, gg, and hh, varying over the inner slope, γ, of the generalized NFW density profile. The bands correspond to γ values spanning 1.2–1.3. Note that the bands for $q\bar{q}$, $c\bar{c}$, and gg fall essentially on top of each other. The best-fit parameters for the $q\bar{q}$, $c\bar{c}$, and gg channels, as obtained in [79], are indicated by the pink, teal, and purple $1\sigma/2\sigma$ filled contours, respectively. The best-fit hh value (and associated 1σ range) is indicated by the blue diamond [79].

astrophysical components because we sum up the pixel-wise likelihoods. To quantify the gain from using spatial templates, we instead perform the fit using only the total expected number of counts from the DM signal and backgrounds within our ROI, and profile over the astrophysical nuisance parameters. The resulting limit (dotted gold line in Fig. 10) is several orders of magnitude weaker than the fiducial bound.

• We show results obtained using the newer `p7v6` and `p8R2` diffuse models in Fig. 11 (left and right panel, respectively). As outlined in the paper, these models have large-scale residuals added back in to various extents, and as such are unsuitable for use in studying large-scale DM structures such as emission from the Galactic halo. Indeed, in both cases, we observe significant over-subtraction for the fiducial ROI. We emphasize that Fig. 11 is included for
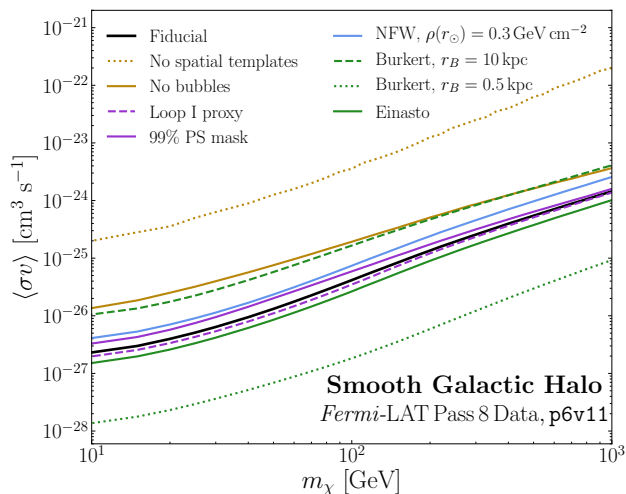
FIG. 10. The 95% confidence limits associated with variations to the fiducial analysis, as labeled in the legend and described in the text.

illustration only and should be treated with caution.

- Because the *Fermi* bubbles are not accounted for when constructing the `p6v11` foreground model, one potential concern is the overestimation of the IC contribution. This could lead to inadequate modeling of the bubbles and potentially over-subtract a DM contribution, leading to an artificially strong limit. We show in Fig. 12 the energy spectra of the Northern (left) and Southern (right) lobes of the *Fermi* bubbles as obtained from our analysis pipeline when using the various foreground models presented here. The spectra recovered when using `p6v11` are broadly similar to those obtained with Models A, B and C, underscoring the fact that the bubbles are adequately modeled in all four cases. We also show the bubbles spectra from [61], obtained for a slightly different ROI ($|b| > 10°$ as opposed to $|b| > 20°$), which are again similar to those derived in our analysis.

- Given the importance of diffuse foreground modeling in the present study and the potential issues associated with a spectrally hard IC component in the `p6v11` model [21], in Fig. 13 we show the total energy spectra obtained for the `p6v11` model as well as those for Models A, B, and C in the eight radial slices considered in our study. We see that the spectra associated with `p6v11` (black line) are roughly consistent with the total spectra associated with Models A, B and C (red, blue and purple lines respectively).

- Figure 14 demonstrates the likelihood profiles for the fiducial analysis. In general, there is very good agreement between the observed profile (black line) and the Monte Carlo expectation (blue band), in each energy bin.
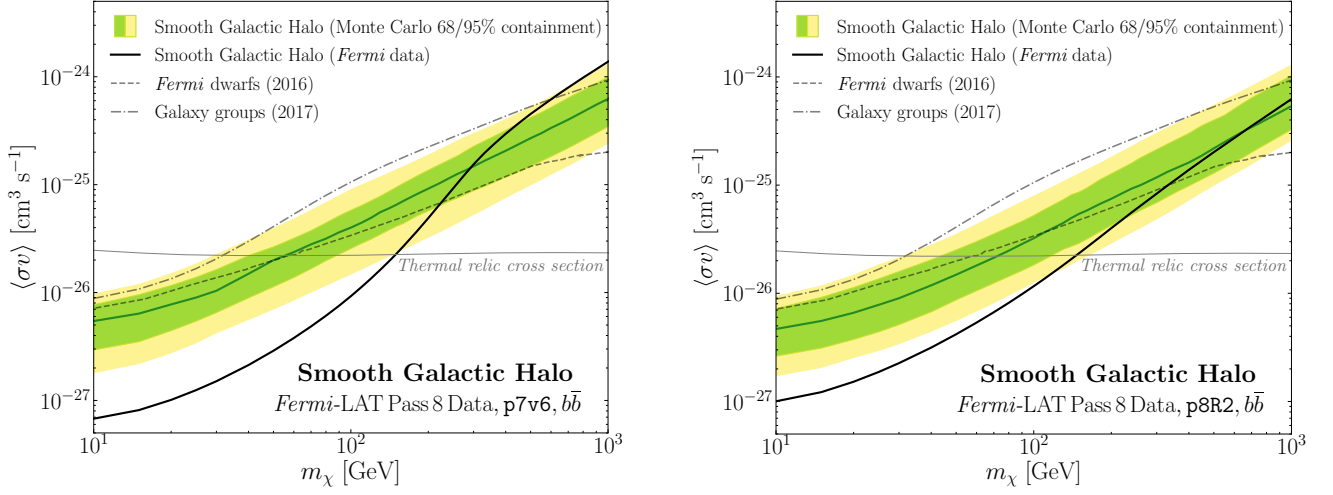
FIG. 11. Similar to Fig. 1 of the paper, except using the `p7v6` and `p8R2` foreground models (left and right panel, respectively) [56]. We only include these results for illustration as both of these foreground models are not appropriate for studies of diffuse DM signals, as discussed in the text.
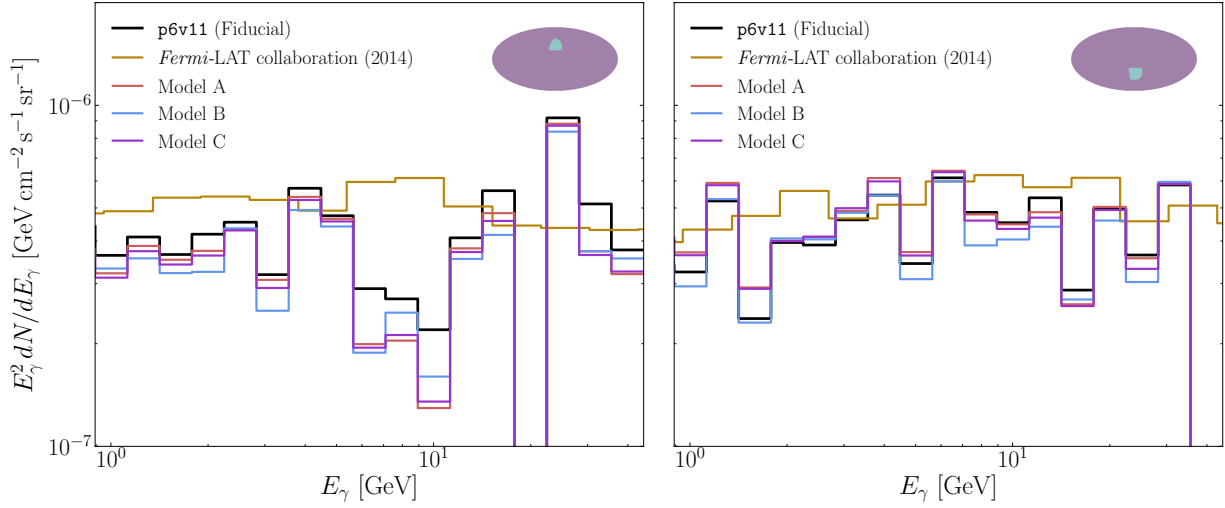

FIG. 12. Recovered spectra, normalized to the corresponding bubbles region shown, for the Northern (left) and Southern (right) lobes of the *Fermi* bubbles when analyzed with diffuse model `p6v11` as well as Models A, B and C. Our fiducial configuration was used to extract these spectra. The bubbles spectra obtained in [61] are shown for comparison. Note that a slighty different ROI ($|b| > 10°$ as opposed to $|b| > 20°$) was used in that case. The energy E_γ corresponds to the geometric mean of the energy bin edges.
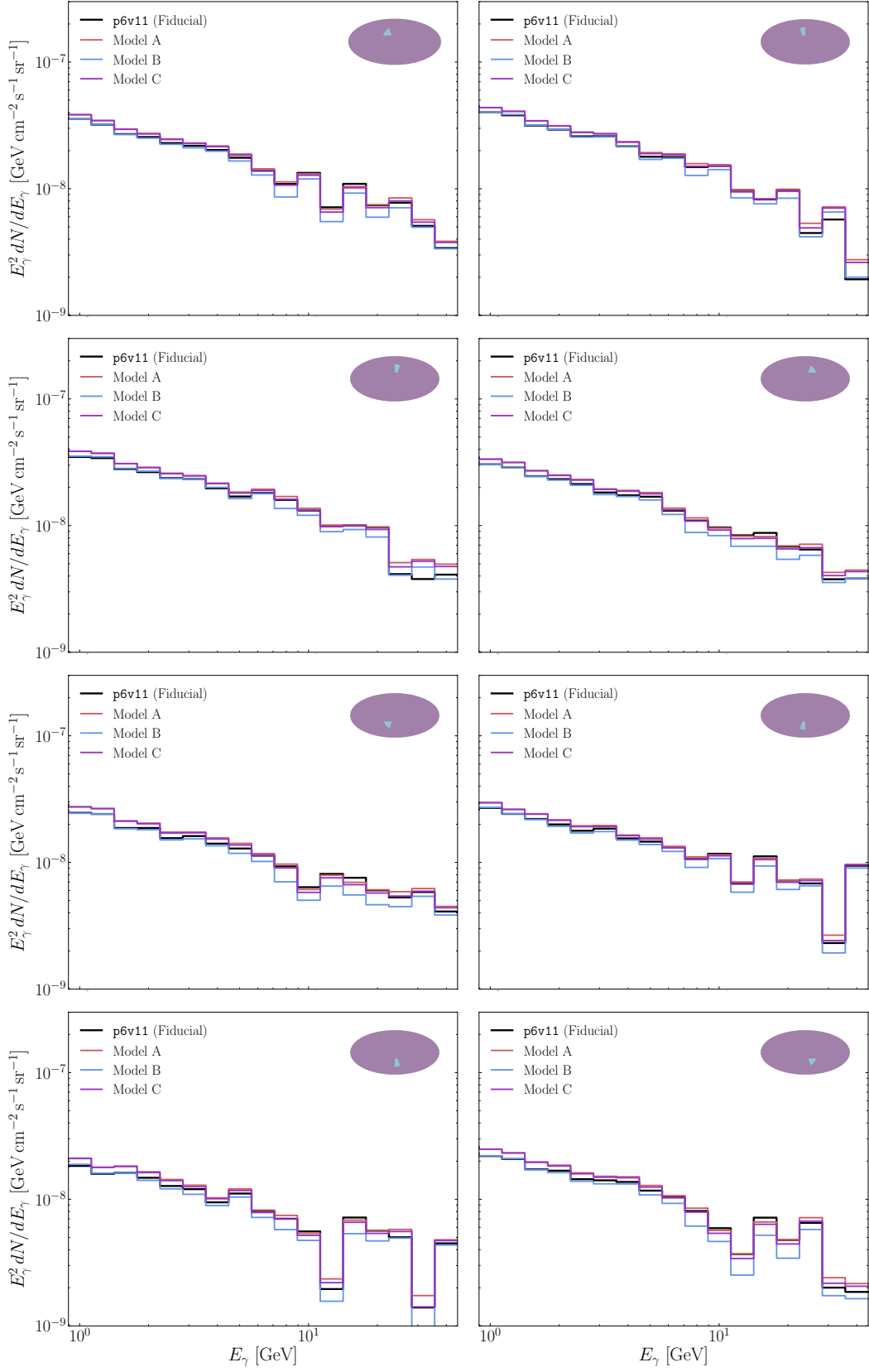
FIG. 13. Energy spectra obtained for the `p6v11` model (black) as well as those for Models A, B, and C (red, blue and purple respectively) in the eight radial slices (shown as insets) considered in our study.
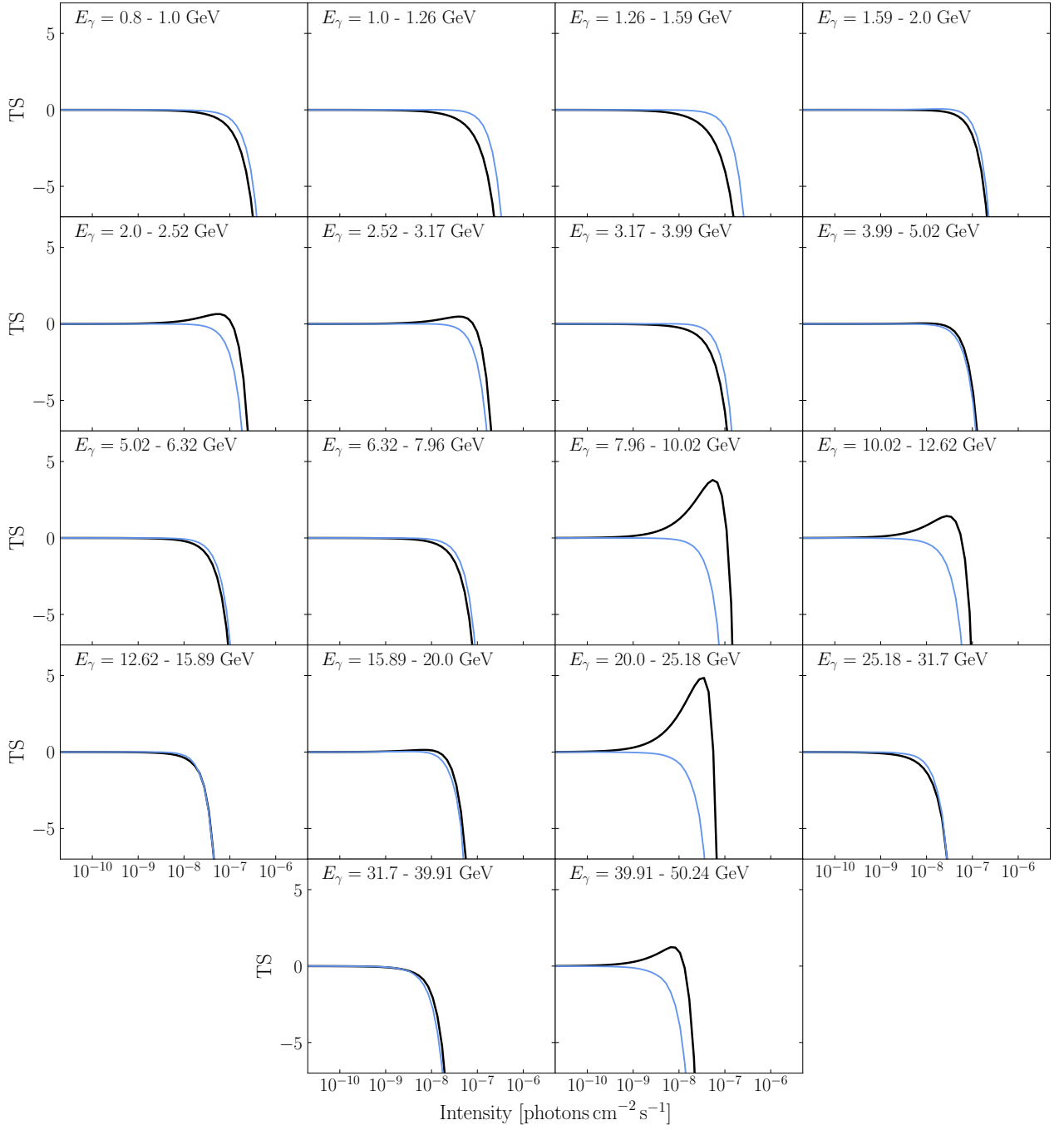
FIG. 14. Likelihood profiles for the fiducial analysis (black lines), presented for each energy bin. The darker(lighter) blue bands denote the 68(95)% containment for the profiles, as determined from 100 Monte Carlo simulations.