# Robust inference of the Galactic Centre gamma-ray excess spatial properties

Deheng Song,[1]★ Christopher Eckner [iD],[2,3,4] Chris Gordon [iD],[5] Francesca Calore,[3] Oscar Macias,[6,7] Kevork N. Abazajian [iD],[8] Shunsaku Horiuchi,[9,10] Manoj Kaplinghat[8] and Martin Pohl[11,12]

[1]*Center for Gravitational Physics and Quantum Information, Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*
[2]*Center for Astrophysics and Cosmology, University of Nova Gorica, Vipavska 11c, 5270 Ajdovščina, Slovenia*
[3]*LAPTh, CNRS, USMB, F-74940 Annecy, France*
[4]*LAPP, CNRS, USMB, F-74940 Annecy, France*
[5]*School of Physical and Chemical Sciences, University of Canterbury, Christchurch 8041, New Zealand*
[6]*Department of Physics and Astronomy, San Francisco State University, San Francisco, CA 94132, USA*
[7]*GRAPPA Institute, University of Amsterdam, NL-1098 XH Amsterdam, the Netherlands*
[8]*Department of Physics and Astronomy, University of California – Irvine, Irvine, CA 92697, USA*
[9]*Center for Neutrino Physics, Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA*
[10]*Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Chiba 277-8583, Japan*
[11]*Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, D-15738 Zeuthen, Germany*
[12]*Institute of Physics and Astronomy, University of Potsdam, D-14476 Potsdam, Germany*

## ABSTRACT

The gamma-ray *Fermi*-LAT Galactic Centre excess (GCE) has puzzled scientists for over 15 yr. Despite ongoing debates about its properties, and especially its spatial distribution, its nature remains elusive. We scrutinize how the estimated spatial morphology of this excess depends on models for the Galactic diffuse emission, focusing particularly on the extent to which the Galactic plane and point sources are masked. Our main aim is to compare a spherically symmetric morphology – potentially arising from the annihilation of dark matter (DM) particles – with a boxy morphology – expected if faint unresolved sources in the Galactic bulge dominate the excess emission. Recent claims favouring a DM-motivated template for the GCE are shown to rely on a specific Galactic bulge template, which performs worse than other templates for the Galactic bulge. We find that a non-parametric model of the Galactic bulge derived from the VISTA Variables in the Via Lactea survey results in a significantly better fit for the GCE than DM-motivated templates. This result is independent of whether a GALPROP-based model or a more non-parametric ring-based model is used to describe the diffuse Galactic emission. This conclusion remains true even when additional freedom is added in the background models, allowing for non-parametric modulation of the model components and substantially improving the fit quality. When adopted, *optimized* background models provide robust results in terms of preference for a boxy bulge morphology for the GCE, regardless of the mask applied to the Galactic plane.

**Key words:** astroparticle physics – pulsars: general – Galaxy: bulge – local interstellar matter – dark matter – gamma-rays: diffuse background.

## 1 INTRODUCTION

The successful deployment of the *Fermi* Gamma-Ray Space Telescope 15 yr ago ushered in an era of unprecedented sensitivity to the gamma-ray sky, with *Fermi*'s Large Area Telescope (LAT) providing increased energy resolution and an unprecedented angular resolution (Atwood et al. 2009). A key science goal of *Fermi*-LAT was to explore gamma-ray emissions from dark matter (DM), specifically investigating the pair annihilation products of thermally produced weakly interacting massive particles (WIMPs).

Shortly after its launch, an observation was reported of an extended source towards the Galactic Centre (GC) consistent with thermal WIMP annihilation with a profile consistent with a cuspy DM halo (Goodenough & Hooper 2009; Vitale & Morselli 2009; Hooper & Goodenough 2011). The presence of excess emission was confirmed and refined with many subsequent studies (e.g. Hooper & Linden 2011; Abazajian & Kaplinghat 2012; Gordon & Macias 2013; Hooper & Slatyer 2013; Abazajian et al. 2014; Calore, Cholis & Weniger 2015; Zhou et al. 2015; Ajello et al. 2016; Daylan et al. 2016; Linden et al. 2016; Ackermann et al. 2017).

This GC excess (GCE) may also be due to an unresolved population of, e.g. millisecond pulsars (MSPs), stellar remnants associated with the central stellar population of the Milky Way (Abazajian 2011; Abazajian & Kaplinghat 2012). However, there

★ E-mail: songdeheng@yukawa.kyoto-u.ac.jp

is debate on whether more GC MSPs should have been resolved (e.g. Hooper & Mohlabeng 2016; Ploeg et al. 2020; Dinsmore & Slatyer 2022; Malyshev 2024) and whether there are sufficient corresponding low-mass X-ray binary (LMXB) systems (e.g. Haggard et al. 2017; Gautam et al. 2022).

One method of distinguishing between the MSP and DM proposals for the GCE is whether the GCE has a spherical morphology (Calore et al. 2015; Ajello et al. 2016; Daylan et al. 2016; Di Mauro 2021; Cholis et al. 2022b; McDermott, Zhong & Cholis 2023) or follows the bulge-like morphology of the GC's stellar populations (Bartels et al. 2018; Macias et al. 2018, 2019; Abazajian et al. 2020; Calore, Donato & Manconi 2021; Pohl et al. 2022). The main obstacle to accurately determining the GCE morphology is uncertainties in the Galactic diffuse emission. When residuals in the gamma-ray fit were significantly reduced either through the construction of advanced gas models (Macias et al. 2018, 2019; Pohl et al. 2022) or through new, more flexible, fitting techniques (Bartels et al. 2018; Calore et al. 2021), a strong preference for the excess tracing old stars in the Galactic bulge emerged. Recently, Di Mauro (2021), Cholis et al. (2022b), and McDermott et al. (2023) have, however, made opposite claims.

Another, independent, way of discriminating between the DM and MSP explanations of the GCE is to look for non-Poissonian statistics in the GCE, which would be indicative of the MSP explanation (Bartels, Krishnamurthy & Weniger 2016; Lee et al. 2016), in contrast to any other truly diffuse signal such as DM. However, this is arguably even more susceptible to uncertainties in the Galactic diffuse emission (Leane & Slatyer 2019; Buschmann et al. 2020; Chang et al. 2020; Leane & Slatyer 2020a,b). New and independent analyses, introducing methodological developments which effectively reduce the impact of Galactic diffuse emission modelling systematics, found a sizable contribution from faint, subthreshold point sources to the GCE (List et al. 2020; Calore et al. 2021; List, Rodd & Lewis 2021; Mishra-Sharma & Cranmer 2022).

Accounting for uncertainties in the Galactic diffuse emission model is undoubtedly key to making robust inferences on the Fermi GCE properties and overcoming the so-called reality gap, i.e. the discrepancy between models and real data (Caron et al. 2023). Different solutions were explored in the literature to overcome systematics related to the Galactic diffuse emission modelling. These approaches either rely on the input from a large number of realizations of cosmic-ray induced diffuse gamma-ray emission as obtained from cosmic-ray propagation codes, like GALPROP[1] (Strong & Moskalenko 1998), (e.g. Calore et al. 2015; Cholis et al. 2022b), or they allow more flexibility in the diffuse-emission model by splitting the Galactic-diffuse-emission templates in multiple rings (e.g. Macias et al. 2018; Di Mauro 2021). Finally, a complementary way to optimize gamma-ray emission model components and, more specifically, the Galactic diffuse emission is the application of data-driven techniques that can reduce the residuals and minimize the gap between model space and reality. A tool that has been utilized in the context of the GCE is SKYFACT (Storm, Weniger & Calore 2017; Bartels et al. 2018; Calore et al. 2021). SKYFACT combines the capabilities of traditional template-based maximum likelihood fits with image reconstruction techniques. Its advantage lies in the addition of spatial and spectral re-modulation for all components of the compiled gamma-ray emission model. In this sense, SKYFACT allows for a more flexible treatment of the employed templates,

especially regarding their spatial morphology, which is fixed in standard template-based fits. To achieve this flexibility, SKYFACT introduces a large number of nuisance parameters and a penalizing likelihood function constraining their variation during the fit in addition to the Poisson likelihood function typically adopted in gamma-ray analyses; for technical details, see Storm et al. (2017).

With this work, we aim at systematically scrutinizing how the morphology of the GCE is affected by background modelling uncertainties when point sources and the disc plane are masked, and analysis approaches are varied, and address in detail some contradictory findings in the recent literature, most notably those in McDermott et al. (2023, M2023 hereafter). As such, our analysis is solely focused on the large-scale emission properties of the GCE and not on its point-source contribution.

In Section 2, we explain the model components and fitting procedure, exploring the background model construction. In Section 3, we focus on template fitting. We show that the ring-based fits of M2023 have not properly converged. We also provide evidence that they have used a bulge template that is inconsistent with the data. We show that by switching to a more accurate bulge template, such as the one generated from the recent VISTA Variables in the Via Lactea (VVV) survey (Coleman et al. 2020), the bulge is favoured over DM-based templates even when M2023's background is used. In Section 3.5, we show that the ring-based methods still find a preference for the Galactic bulge over the DM template even when the point sources masks are substantially increased in size. We also show that after masking, the inclusion of a Galactic bulge passes a Monte Carlo based goodness of fit test. In Section 4, we employ a skyFACT modulation of the diffuse templates (Storm et al. 2017; Bartels et al. 2018; Calore et al. 2021). We find that in all cases, the Bayesian evidence is improved by the modulation, and the addition of the DM template is not favoured once the modulation has been done and a Galactic bulge template has been included. The conclusions are given in Section 5.

## 2 GAMMA-RAY MODEL COMPONENTS AND FITTING PROCEDURE

In this section, we first provide a brief overview of the components used to model the gamma-ray sky. In general, the inner Galaxy sky is interpreted as the sum of the following main contributions:

(i) Galactic diffuse emission: This contribution, which we will discuss in more detail below, is the result of the interactions of cosmic rays with the interstellar gas and low-energy interstellar radiation fields. The main contributors to the Galactic diffuse emission are the decay of neutral pions produced in collisions between cosmic-ray protons and interstellar gas, the inverse Compton scattering of the interstellar radiation field by electrons, and the bremsstrahlung emission from these electrons.

(ii) Point-like and extended sources: These correspond to gamma-ray identified sources which are listed in the *Fermi*-LAT catalogues and can be masked or refitted in the analysis. We will discuss the treatment of point-like sources in the following sections.

(iii) Isotropic diffuse gamma-ray background emission (IGRB): This component accounts for the (almost) isotropic emission measured at high latitudes and is thought to originate from the superposition of different contributions mostly from extragalactic, faint sources (Ackermann et al. 2015).

(iv) Galactic Centre excess, GCE: To complete the description of the inner Galaxy gamma rays, it has been demonstrated that one needs to consider an additional contribution, the so-called GCE.

---

[1] http://galprop.stanford.edu

## 2.1 The Galactic diffuse emission models

Given the vast number of independent degrees of freedom across the sky, constructing an expected Galactic diffuse emission model requires numerous assumptions. Works like Cholis et al. (2022b, hereafter C2022) postulate that the inner galaxy is predominantly influenced by Galactic diffuse emission stemming from nearly steady-state astrophysical processes and modelled through standard propagation codes like GALPROP. In the following, we will refer to these models as 'GALPROP-based' templates. In contrast, other approaches, such as those in Macias et al. (2018, 2019), Abazajian et al. (2020), Pohl et al. (2022), introduce concentric cylindrical, galactocentric-based templates, which we will refer to as 'ring-based templates'.

There are conflicting reports in the literature regarding the best method to use for modelling the Galactic diffuse emission. First, on the background model construction, Pohl et al. (2022, hereafter P2022[2]) demonstrated that the ring-based method provided a better fit to the *Fermi*-LAT data than the non-ring-based method of Di Mauro (2021). This was followed by M2023, who reported that when the point sources and Galactic plane were masked, the GALPROP-based templates, generated by C2022, provided a better fit to the *Fermi*-LAT data. Secondly, there are conflicting reports of whether the GCE is spherical DM-like or bulge-like. On one hand, P2022 found that bulge models were better explanations of the data, while, on the other hand, M2023 claimed a better fit for the spherical DM template as opposed to the Galactic bulge.

In this work, we systematically compare 'GALPROP-based' and 'ring-based' templates for the Galactic diffuse emission, with the final goal of better understanding some contradictory claims in the literature.

In summary, the templates that compose the Galactic diffuse emission in the two approaches are as follows:

**'GALPROP-based' model:** C2022 and M2023 made use of GALPROP to propagate cosmic rays within the interstellar medium and evaluated the diffuse emission templates. The diffuse emission of pion decay is combined with the subdominant bremsstrahlung contribution into one template. Inverse-Compton (ICS) emission was also included. The templates from GALPROP were augmented by templates for the isotropic background and Fermi bubbles. In total, the 'GALPROP-based' model has four templates.

'Ring-based' model: As in P2022 and M2023, the ring-based approach utilizes 16 independent galactocentric cylinders. This model comprises four rings for the neutral atomic hydrogen (H I) density and another four for the molecular hydrogen (H2) density. Additionally, six rings follow the ICS emission. Also included are two dust-derived templates, one negative and one positive valued. The positive dust correction template physically represents H I and H2 hydrogen that is not traced by the relevant emission, known as the dark neutral medium, or an overestimation of the atomic hydrogen spin temperature (Acero et al. 2016). The negative dust correction template represents an underestimation of the spin temperature. The H I and H2 rings, shaped as annular cylinders, have boundaries located at 3.5, 8, 10, and 50 kpc from the GC. The ICS rings share these boundaries, except the innermost ring is further subdivided at radii of 1.5, 2.5, and 3.5 kpc. Also incorporated are the identical isotropic background and Fermi bubbles template as was done by M2023. A total of 18 templates are used for the 'ring-based' model.

## 2.2 The Galactic Centre excess templates

As done by the majority of the literature, in this work we will consider two main hypotheses for the morphology of the GCE.

The first one is that the signal is connected with annihilating DM in the inner Galaxy. To this end, we consider DM-inspired templates constructed as a spherical template that follows the square of a generalized Navarro–Frenk–White profile (gNFW[2]):

$$\rho(r) \propto \frac{1}{(r/r_c)^\gamma \, (1 + r/r_c)^{3-\gamma}}. \tag{1}$$

We adopt the parameters used in C2022 and M2023, which are $\gamma = 1.2$ and $r_c = 20$ kpc. The DM template is then obtained by integrating $\rho(r)^2$ along each line of sight. Note, however, that there is currently no clear prediction for the inner (within about 2 kpc) density profile of the DM halo of the MW (or its sphericity) with both cusps and cores found in sophisticated hydrodynamic simulations (Lazar et al. 2020; Grand & White 2022).

The second hypothesis builds on the fact that the GCE may trace the distribution of old stars in the Galactic bulge. Direct observations of the GC region have historically been obscured due to challenges posed by dust reddening. The advent of near-infrared surveys, such as the ground-breaking COBE/DIRBE study, has enabled significant advancements in our understanding of this central region (Bland-Hawthorn & Gerhard 2016). These surveys have not only confirmed the presence of the Galactic bulge/bar but also provided the foundational data upon which subsequent triaxial bar models of our galaxy were constructed (Binney et al. 1991; Weiland et al. 1994; Freudenreich 1998). Cao et al. (2013) utilized red clump giants from the OGLE-III survey to create a detailed photometric model of the Galactic bar. This relied on data available up to 2013. In an advancement, Coleman et al. (2020) utilized more contemporary data from the VVV survey. Their approach, integrating non-parametric methodologies, offers a flexible means of estimating the morphology of the Galactic bulge, thus refining our understanding of this critical Galactic structure. The flexibility is important because we want the data and not an assumed inflexible functional form, as in Cao et al. (2013) for example, to determine the radial profile of the template.

In this work, we consider several different bulge templates. The bulge model used in M2023 is publicly available as part of the analysis package GCEPY[3] Following M2023, we will label this the boxy-bulge BB (GCEPY) template. In addition to one from GCEPY, we consider three other bulge models: Freudenreich (1998, hereinafter F98), Cao et al. (2013, hereinafter Cao13), and Coleman et al. (2020, hereinafter Coleman20). We notice that the BB template in GCEPY appear be obtained from the Cao13 model (M2023, personal communication), but upon further inspection, we found that it does not match our version of the Cao13 template. We therefore keep the BB GCEPY template and the Cao13 as different, independent, choices for the bulge. Fig. 1 shows the spatial templates for these bulge models in the main region of interest (ROI) of this work, i.e. $40° \times 40°$ around the GC. We have normalized the templates to have the same flux in the ROI (in an arbitrary unit). The contours on the maps show the 10 per cent, 30 per cent, 50 per cent, 70 per cent, and 90 per cent levels with respect to the central value. The contours demonstrate the differences in the bulge models, both at large scales as well as near the centre.

On top of the boxy-bulge, we also add a component for the nuclear bulge (NB) following the parametric model of Launhardt, Zylka &

---

[2]Note that there is some overlap in the authors of the current article and those of P2022.
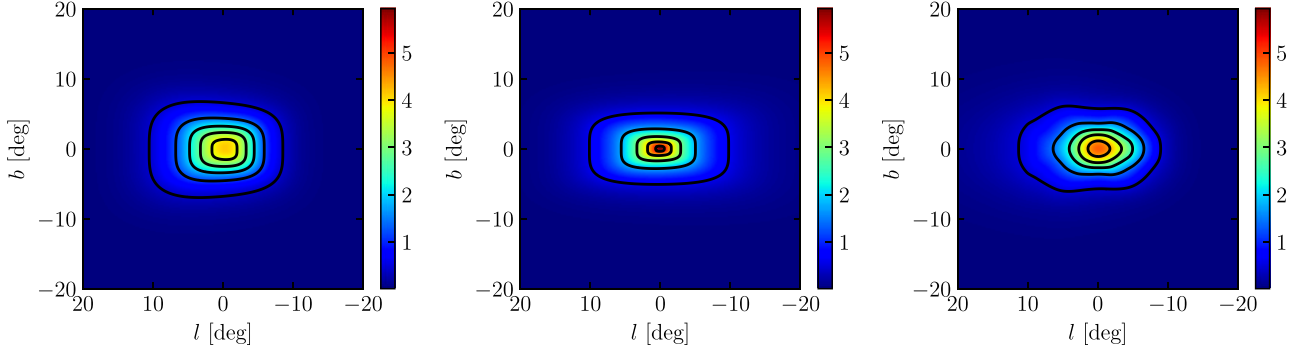
[3]https://github.com/samueldmcdermott/gcepy

**Figure 1.** Spatial templates of the Galactic bulge models considered in this work. References from left to right: Freudenreich ([1998]), Cao et al. ([2013]), and Coleman et al. ([2020]). Note that these are line-of-sight integrated images of the bulge templates before they are convolved with the PSF.

Mezger ([2002]), which is necessary to include when not masking the Galactic plane (see Sections 3.5 and 4).

### 2.3 Fitting procedure and statistical framework

In this paper, we adopt two different fitting techniques. The first one is the traditional 'template fit', where all model components are defined as morphological templates that are not allowed to vary during the fit procedure. In template fitting, the free parameters are the independent bin-by-bin normalizations of the spectral energy distribution of each model component. The second technique is the so-called adaptive template fitting, where the spatial distribution of photons among different sky components is enabled to be re-modulated during the fitting procedure (Storm et al. [2017]; Bartels et al. [2018]; Calore et al. [2021]). In what follows, we briefly introduce more technical details about the implementation of both data fitting approaches. We conclude this section by outlining the statistical inference methods that we apply to investigate the morphology of the GCE.

#### 2.3.1 Traditional template fitting

In standard template fit analysis, the gamma-ray sky is described as the sum of multiple model components, $k$, identified by a fixed spatial distribution, i.e. the spatial template. The model $\phi$ is a linear combination of $k$ spatial templates $T$ binned in spatial pixels $p$ and a spectral normalization $\Sigma$ for each energy bin $b$, which is free to re-adjust during the fit:

$$\phi_{pb} = \sum_k T_p^{(k)} \cdot \Sigma_b^{(k)}. \tag{2}$$

The optimization of the free energy-independent normalization is done by maximizing the Poissonian likelihood given the number count maps.

For the optimization, M2023 employs DYNESTY to scan the parameter space and then uses a No-U-Turn sampler (utilizing NUMPYRO) to optimize the best-fitting result. For the sake of cross-checking the M2023 results, in Section 3, we will perform the maximum likelihood analysis with an alternative minimizer, i.e. using the Limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm extended to handle simple box constraints (L-BFGS-B). For implementation, we use the PYTHON package LMFIT. We use `stats.poisson.logpmf` provided by SCIPY to calculate the Poissonian log-likelihood. Following M2023, we also include penalty terms (provided by GCEPY) on the log-likelihood to account for when the IGRB and *Fermi* bubbles

normalizations deviate too much from their spectra measured at high latitudes.

Point sources in the 4FGL are masked for this analysis. We will provide more details on the mask implementation in Section 3.

#### 2.3.2 Adaptive-constrained template fitting

A complementary analysis technique to investigate the preferred morphology of the GCE is adaptive template fitting implemented in the software package SKYFACT. As mentioned in the introduction, SKYFACT is a mixture of a conventional template-based fit and more advanced image reconstruction routines. The main advantage is a departure from imposing a non-flexible prior on the spatial morphology of the compiled background and signal components of the gamma-ray emission model. While this flexibility enables capturing and remedying a certain degree of component mis-modelling, it requires the introduction of a large number of nuisance parameters that have to be controlled during the fit to avoid overfitting and unphysical results. Therefore, SKYFCAT utilizes a combination of a Poisson and a penalizing likelihood function to guide the fit with constrained freedom for all nuisance parameters. The fit itself is the minimization of the mentioned log-likelihood function.

In essence, the gamma-ray emission model compiled for SKYFACT is identical to the input data required for a traditional template fit. That is, the model $\phi$ is a (tri-)linear combination of $k$ spatial templates $T$ binned in spatial pixels $p$ and a spectral normalization $S$ per energy bin $b$ following a certain functional form or tabulated data:

$$\phi_{pb} = \sum_k \tau_p^{(k)} T_p^{(k)} \cdot \sigma_b^{(k)} S_b^{(k)} \cdot \nu^{(k)}. \tag{3}$$

Examples of suitable spatial templates are the Galactic bulge models in Fig. 1. The addition of SKYFACT is to introduce a global normalization parameter $\nu$ per component and spatial and spectral nuisance parameters, $\tau$ and $\sigma$, respectively. In practice, $\tau$ and $\sigma$ act as a one-to-one copy in shape of the spatial and spectral input priors. Both nuisance parameter sets are initialized with value 1 and varied – or better, *re-modulated* – in the priors' stead during the fitting procedure. The nuisance parameters are required to be strictly positive.

We note that the SKYFACT model $\phi$ allows for the inclusion of point-like sources according to their position and spectra listed in gamma-ray catalogues of choice. For them, the position inside the considered region of interest is fixed, i.e. there is no associated spatial template $T$. Consequently, all injected point-like sources are spectrally re-fit in adaptive template fitting.

The numerous nuisance parameters are tamed via the regularizing part of SKYFACT's likelihood function, which depends on five hyper-parameters per component[4] constraining the magnitude of individual parameters and the correlation among neighbouring parameters. All details about the explicit structure of the likelihood function and hyper-parameters is provided in Storm et al. (2017). Here, it shall suffice to give an illustrative example of the functionality of two of them, the spatial and spectral smoothing scale $\eta$ and $\lambda$. They are defined as $\eta = 1/x^2$ ($\lambda = 1/E^2$), where $x$ ($E$) is the allowed variation between neighbouring spatial pixels (energy bins). Setting the smoothing scales to zero is thus equivalent to saying that all spatial pixels (energy bins) may vary entirely independently of each other.

As the tuning of the hyperparameters is cumbersome due to computational speed, we select the configuration of RUN5 in Storm et al. (2017) as the foundation for our settings. The specifications of RUN5 assume a certain spatial smoothing for the brightest gamma-ray components, namely the inverse Compton component receives the largest smoothing, followed by the $\pi^0$-component and lastly the *Fermi* bubbles with the shortest smoothing scale. Regarding the remaining model elements we do not allow for any spatial re-modulation of the input template, this applies to the different GCE components, in particular. We hence ensure that the minimization problem remains convex guaranteeing convergence of the L-BFGS-B algorithm utilized in SKYFACT. Spectral smoothing is not applied at all.

### 2.3.3 Statistical framework: parameter inference and model comparison

In a frequentist setting, it is possible to properly quantify the preference for different gamma-ray emission models based on the maximum likelihood method and the resulting likelihood values at the optimal point. However, a well-defined notion of model comparison from a statistical point of view is limited to nested models. By nested models we mean any gamma-ray emission model that is comprised of a default set of components (or base components) and to which further components are added without removing any contributions from the default set. Then it is possible to compute the preference in the data for the enlarged model over the base model in terms of significance.

To be more quantitative let us assume that we add a single component $X$ to the base model and perform a standard template fit so that the extended model has $N$ additional parameters (normalizations per energy bin). Let $\ln \mathcal{L}_{\mathrm{base}}$ denote the log-likelihood value evaluated for the best-fitting parameters determined via the maximum likelihood approach. Likewise, $\ln \mathcal{L}_{\mathrm{base+X}}$ is the corresponding log-likelihood value found for the extended base model including $X$. We choose the log-likelihood ratio test statistic as a means to quantify the significance: $\mathrm{TS} = 2(\ln \mathcal{L}_{\mathrm{base+X}} - \ln \mathcal{L}_{\mathrm{base}})$. Within this set-up, the test statistic is distributed according to a mixture distribution following (Macias et al. 2018)

$$p(\mathrm{TS}) = 2^{-N}\left[ \delta(TS) + \sum_{k=1}^{N} \binom{N_{\mathrm{dof}}}{k} \chi_k^2(\mathrm{TS}) \right]. \tag{4}$$

Here, $\delta$ refers to the Dirac distribution, $\binom{n}{k}$ is the binomial coefficient, and $\chi_k^2$ denotes a $\chi^2$-distribution with $k$ degrees of freedom. The

---

[4]Only three in the case of point-like sources.

significance $\sigma$ of the added component under the observation of the test statistic value $\hat{\mathrm{TS}}$ amounts to

$$\sigma(\hat{\mathrm{TS}}) = \sqrt{\mathrm{CDF}^{-1}\left( \chi_1^2, \mathrm{CDF}\left( p(\mathrm{TS}), \hat{\mathrm{TS}} \right) \right)}, \tag{5}$$

where $\mathrm{CDF}(f, x)$ refers to the cumulative distribution function of $f$ at $x$ and $\mathrm{CDF}^{-1}$ is its inverse.

Model comparison is challenging to correctly perform in a frequentist framework as shown above, while it is better defined and easier to access in a Bayesian approach. We also adopt the latter to run model comparison. Deriving the Bayesian evidence of each individual gamma-ray emission model allows us to perform a Bayesian model comparison or hypothesis testing based on the Bayes factor. Given the Bayesian evidence $\ln \mathcal{H}_X$ of model X and $\ln \mathcal{H}_Y$ of model Y, the mutual Bayes factor is given by $\mathcal{B}_{XY} = \exp(\ln \mathcal{H}_X - \ln \mathcal{H}_Y)$. A positive value of the Bayes factor implies a certain degree of evidence for model X being preferred over model Y. In what follows, we utilize the empirical classification of the degree of evidence from table 1 of Trotta (2008) based on the logarithmic Bayes factor.[5]

While the Bayesian framework offers a direct way of stating the preference for one gamma-ray emission model over another one – irrespective of the exact composition – it depends by construction on prior probabilities for all parameters. In this sense, it carries a certain intrinsic user bias due to the choice of priors, may it be their parametric shape or range. Consequently, we remark as a caveat that the derived preference for a model is prior-dependent. As we will show later in Section 3.3, a suitable choice for the prior range is essential to cover the correct best-fitting point in the model's parameter space.

## 3 TEMPLATE FITTING: REPRODUCIBILITY OF PREVIOUS WORKS AND IMPROVEMENTS

In this section, we weigh in on the findings of M2023, first reproducing and then improving on their analysis. The fitting technique here adopted is the traditional template fit (Section 2.3). We will perform Bayesian model comparison to assess what is the best model for describing gamma-ray data among the ones we test. We will also discuss how the evidence for the DM-inspired template is affected by different choices of point sources and Galactic plane mask, by making use of nested models in the frequentist approach.

### 3.1 Reproducing M2023

As a first step, we repeat the analysis performed in M2023, by adopting the same data set and models. We remind the reader that the M2023 analysis was performed using both GALPROP- and ring-based background models with standard template fitting, and the authors claimed that the GCE is better described by a DM-like model than a bulge one. A summary of the data selection in C2022 and M2023 is reported in Table 1. We adopt this data set as publicly available through the GCEPY webpage. Together with the selected data set (counts map), the GCEPY package released also (i) Galactic plane and point-source mask adopted, and (ii) the model templates convolved with the *Fermi*-LAT point spread function (PSF). The mask adopted by M2023 (and in this section unless otherwise specified) masks

---

[5]We also provide the definition of TS between different models, as $TS = 2(\ln \mathcal{L}_X - \ln \mathcal{L}_Y)$. While we refrain from assessing model comparison through this $TS$, we notice however that this practice is largely present in the previous literature. Therefore, for the sake of comparison, in the main results' tables we will also report this value without overinterpreting it.

**Table 1.** Data selection according to C2022 and M2023.

| Parameter | Value |
|---|---|
| Time range | Week 9 to Week 670 (12.5 yr) |
| Energy range | 14 bins from 0.275 to 51.9 GeV |
| ROI | $40° \times 40°$, bin size $= 0.1°$ |
| Data class | P8R3_CLEAN_V3 (evclass = 256), FRONT (evtype = 1) |
| Filter | (DATA_QUAL==1)&&(LAT_CONFIG== 1) |
|  | &&(ABS(ROCK_ANGLE)<52) |
| Max zenith angle | 100 deg |

out both the Galactic plane ($|b| < 2°$) and the point sources in the 4FGL-DR2 catalogue.

In M2023, two GALPROP-based models are identified as best-fitting ones: GALPROP$_{7p}$ is, according to M2023, the best Galactic diffuse model when no GCE is added to the fit; GALPROP$_{8t}$ instead provides the best performance when adding a GCE modelled through a gNFW$^2$ profile. In C2022's Zenodo archive (Cholis et al. 2022a), these two cases can be seen to correspond to the cases XLIII and XLIX, respectively, in table VIII of C2022.

We run a traditional template fit though maximum likelihood optimization, as described in Section 2.3. The minimization is run with both the GCEPY code and our implementation with the L-BFGS-B algorithm.

The results from the runs reproducing M2023 are reported in Table 2. Following M2023, we use the background model GALPROP$_{8t}$, corresponding to the best-fitting model obtained when an additional gNFW$^2$ component has been added. As done by M2023, we compare this to the best-fitting background in the case of no GCE source (GALPROP$_{7p}$). Compared with the scenario without GCE, the BB GCEPY template has a $\ln \mathcal{B} = 885$, indicating evidence for a GCE. We remind the reader that no NB is included in the model, following M2023.

For GALPROP-based background models, we find very similar likelihood values no matter what is the adopted minimization procedure. This contrasts to what happens for ring-based templates (see below).

## 3.2 Model systematics I: testing other bulge templates

We repeat the analysis performed in M2023, but with additional bulge models described in Section 2. To consistently convolve with the *Fermi*-LAT PSF (which are not included in the original M2023), we run FERMITOOLS based on the same criteria as in C2022 and M2023 (see Table 1 for details). Namely, we use GTSELECT and GTMKTIME to select and filter the events, then use GTBIN to bin the data. After using GTLTCUBE and GTEXPCUBE2 to obtain the live-time and exposure, we use GTSRCMAPS and GTMODEL to generate the convolved templates of the three bulge models. In Fig. 2, we display the three additional bulge templates after they are convolved with the PSF at 1.02–1.32 GeV, compared with the M2023 one.

Table 2 reports the results using the GALPROP-based background model also for the different bulge templates. The mutual Bayes factor, $\ln \mathcal{B}_{XY} \equiv \Delta \ln \mathcal{H}$, allows us to appreciate the performance of the different models and to assess which one performs better. From Table 2, we can see that our version of the Cao13 template has $\ln \mathcal{B} = 180$, when compared to the BB GCEPY model, meaning that it provides a better model for the gamma-ray data. On the other hand, our other two bulge models are even better than the BB GCEPY template, with $\ln \mathcal{B} = 507$ and 685 for F98 and Coleman20, respectively, when compared to the BB gcepy model.

As for the comparison with the DM-inspired template, we see that the gNFW$^2$ is preferred over the BB GCEPY and Cao13 model, consistent with the findings of M2023. However, F98 has $\ln \mathcal{B} = 54$, when compared to gNFW$^2$. Finally, the Coleman20 model results to be the best GCE template in our tests for the GALPROP-based background model, with a $\ln \mathcal{B} = 232$, when compared to gNFW$^2$. We conclude, therefore, that for GALPROP-based models, the choice of the bulge template is crucial for interpreting whether the GCE is DM like or bulge like. We nonetheless stress that the above statement holds for the specific diffuse emission model adopted, which is not guaranteed to provide an overall good fit to gamma-ray data. We will discuss how ring-based models improve the goodness of fit in the next section.

## 3.3 Model systematics II: testing ring-based models

We here explore the same set of four bulge models but with the ring-based Galactic diffuse models of P2022. We remind the reader that M2023 also presented a run with the P2022 ring-based background model, arguing that this always provided a worse fit to the data than the GALPROP-based optimized background model when no excess was considered, i.e. GALPROP$_{7p}$.

Also in this case, we run the fit with both minimizers to cross-check the validity of the GCEPY code. Differently from the GALPROP-based models, when comparing the results of the two minimizers for the ring-based background model, we find a major discrepancy between our results and M2023: Contrary to the finding of M2023, when using the L-BFGS-B algorithm we find that the ring-based background model provides a better fit compared with the GALPROP$_{7p}$ background model.

In Table 3, we compare the likelihood results[6] from the L-BFGS-B algorithm and the GCEPY package. We focus on the no-excess case and compare the GALPROP-based and ring-based background models. Using the L-BFGS-B algorithm, we find that the TS for the GALPROP-based background model against the ring-based one is $-1852$. Using the GCEPY package, we find instead that the GALPROP-based background model has a positive TS of 3508 against the ring-based one. This observation aligns qualitatively with the findings in M2023 who report a positive TS of 4539 for the GALPROP-based background. Table 3 shows that the $-2 \ln \mathcal{L}$ values for the GALPROP-based background model are almost the same between L-BFGS-B and GCEPY. However, this is not the case for the ring-based background model. Thus, the discrepancy is only seen in the ring-based analyses.

After investigating the fitting results in each energy bin and for every template, we find that the major difference between the fits using the L-BFGS-B algorithm and GCEPY is in the best-fitting values for the negative and positive dust corrections in the ring-based model. To use nested sampling to estimate Bayesian posteriors, GCEPY has to implement priors for the templates. The adopted priors are uniform in logarithmic space and are sufficiently wide for most templates. However, the priors for the negative and positive dust corrections turned out to be too limiting. In the public code of GCEPY, the priors for the normalization in logarithmic space are uniform between $[-2, 4]$ for the negative dust correction and $[-2, 6]$ for the positive dust correction. However, when we adopted the L-BFGS-B algorithm, we

---

[6] The L-BFGS-B algorithm simply optimizes the likelihood and therefore does not allow for Bayesian evidence calculation. For this comparison, we, therefore, use $\Delta$TS values, but we will show that our conclusions also hold when considering the Bayesian evidence.

**Table 2.** Results of the likelihood analysis for the GALPROP-based model using GCEPY, including the Bayesian evidence, and mutual Bayes factor with respect to the best-fitting model GALPROP$_{7p}$. We remind that the same mask as in M2023 is applied on 4FGL-DR2 sources and the Galactic plane, $|b| < 2°$. Likelihood values from GCEPY are consistent with the ones obtained with the L-BFGS-B algorithm.

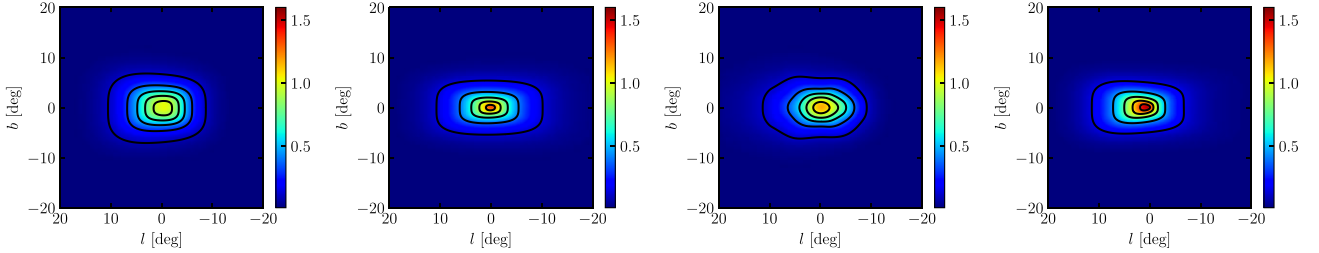| Baseline model | Additional source | $-2\ln\mathcal{L}$ | TS | $\ln\mathcal{H}$ | $\ln\mathcal{B} \equiv \Delta\ln\mathcal{H}$ |
|---|---|---|---|---|---|
| GALPROP$_{7p}$ | none | 3 752 798 | 0 | −1876678 | 0 |
| GALPROP$_{8t}$ | BB (GCEPY) | 3 750 941 | 1857 | −1875793 | 885 |
| GALPROP$_{8t}$ | gNFW$^2$ | 3 750 051 | 2747 | −1875340 | 1338 |
| GALPROP$_{8t}$ | Cao13 | 3 750 582 | 2216 | −1875613 | 1065 |
| GALPROP$_{8t}$ | F98 | 3 749 924 | 2874 | −1875286 | 1392 |
| GALPROP$_{8t}$ | Coleman20 | 3 749 563 | 3235 | −1875108 | 1570 |



**Figure 2.** Bulge templates after they are convolved with the PSF at 1.02–1.32 GeV. References from left to right: Freudenreich (1998), Cao et al. (2013), Coleman et al. (2020), and M2023.

**Table 3.** Comparison between the L-BFGS-B algorithm and GCEPY with its original priors for the GALPROP$_{7p}$-based background model and the ring-based background model without GCE.

| Background model | Fitting algorithm | $-2\ln\mathcal{L}$ | $2\left(\ln\mathcal{L}_{\mathrm{GALPROP_{7p}}} - \ln\mathcal{L}_{\mathrm{ring-based}}\right)$ |
|---|---|---|---|
| GALPROP$_{7p}$ | L-BFGS-B | 3 752 792 | −1852 |
| Ring based | | 3 750 940 | |
| GALPROP$_{7p}$ | GCEPY | 3 752 798 | +3508 |
| Ring based | | 3 756 306 | |

provided bounds to scale the parameters from 0 to large values, far exceeding the priors in GCEPY in linear space. Fig. 3 shows the best-fitting values for the normalization of the negative and positive dust corrections from L-BFGS-B and GCEPY in each energy bin for the ring-based background model. It is clear that the best-fit values for the negative dust correction from L-BFGS-B always exceed the prior range of GCEPY. In GCEPY, the best fit found by GCEPY simply stops at the upper boundary of the prior for most bins. For a few high-energy bins, the best fit found by GCEPY is very small, likely caused by finding a local minimum. The same situation is also observed for the positive dust correction, although only for a few high-energy bins. In Table 4, we find that the best-fitting likelihood for the ring-based model from GCEPY is again consistent with that from the L-BFGS-B algorithm once we widen the priors for dust corrections. More specifically, we widen the prior upper bound of two dust corrections to 10 while maintaining the other priors unchanged. We conclude that M2023 failed to find the real best-fitting models when using the ring-based background model due to inadequate priors for the dust corrections. Our results using the L-BFGS-B algorithm in Table 3, and the GCEPY results with wider priors in Table 4, provide a more accurate interpretation of the ring-based background model.

In the 'no-excess' case, as detailed in Fig. 4, the best-fitting spectra include the four H I and H2 rings and six ICS rings, along with both negative and positive dust corrections. Notably, the negative dust correction at GeV energies is about 30 per cent of the total HI and H2 fluxes, while the positive correction is relatively small. As the

negative dust correction template represents an underestimation of the spin temperature, we would expect a constant ratio of negative dust correction to the H I spectrum across all energies. Although, this appears to mainly be the case, a deviation occurs at the highest energy bin. However, this discrepancy is not crucial for determining the GCE's morphology, as the GCE's values are minimal at such high energies.

Fig. 5 displays the best-fitting count map (in $\log_{10}$ scale) for the gas-correlated component (H I, H2, and the dust corrections) in the ring-based background model at the 1.02–1.32 GeV energy bin. For comparison, we also show the gas-correlated component (pion decay + bremsstrahlung) in the best-fitting GALPROP-based background model for the same energy bin. Due to the negative correction, the photon counts associated with the gas-correlated component in the ring-based model are generally lower than those in the GALPROP-based model. However, no pixels in the unmasked region have negative values when H I, H2, and dust corrections are combined. We have verified that this holds true for every energy bin.

In Table 4, we add the results from using the ring-based background model using wider priors in GCEPY. In this case, again, likelihood values with the alternative minimizer are comparable. Comparing row one of Table 2 with row one of Table 4 shows that the ring-based model without GCE is a better description of the gamma-ray sky than the GALPROP$_{7p}$, with a $\ln\mathcal{B} = 216$. When adding a GCE component, regardless of the choice of the template, the Bayesian evidence for the GCE is overall reduced for
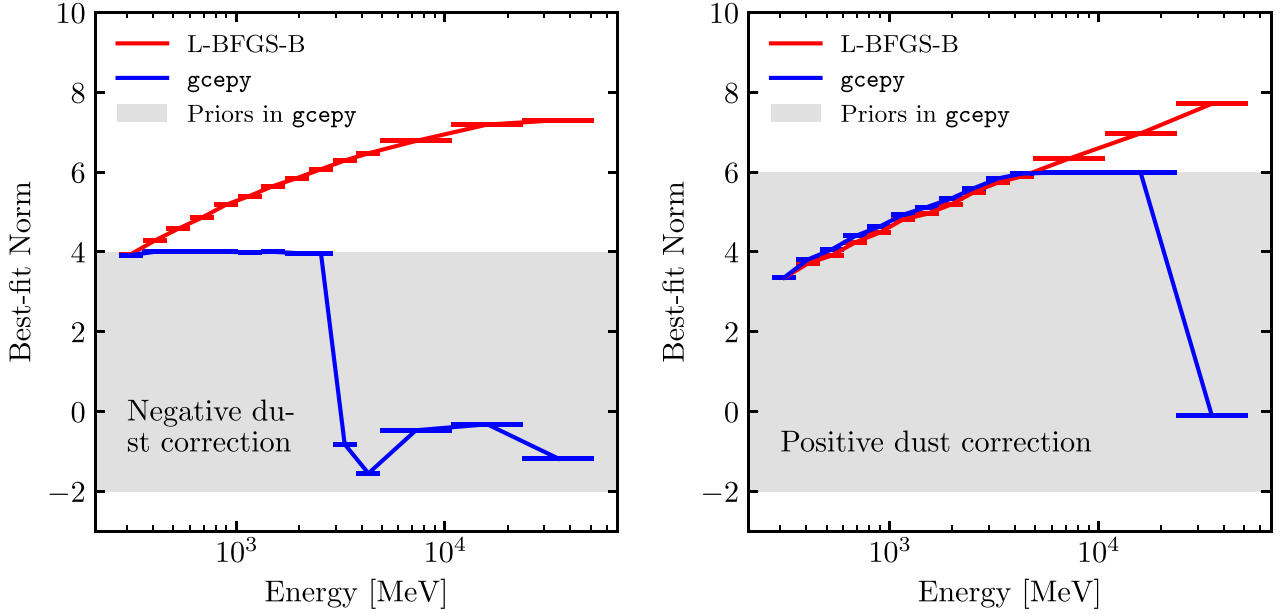
**Figure 3.** Best-fitting normalizations for the negative and positive dust corrections from L-BFGS-B and GCEPY (with narrow priors as in the original implementation) using the ring-based background model developed by P2022. We are plotting the dust template normalization against the 14 energy bins range from 0.275 to 51.9 GeV.

**Table 4.** Similar to Table 2 except here we consider the ring-based background model developed by P2022. We have expanded the priors for the dust corrections in GCEPY to ensure convergence. Note that in each row, the amplitude of the ring-based background model templates is optimized in addition to the GCE additional source (if there is one) to maximize the likelihood $\mathcal{L}$, as explained in Section 2.

| Baseline model | Additional source | $-2\ln\mathcal{L}$ | TS | $\ln\mathcal{H}$ | $\ln\mathcal{B} \equiv \Delta\ln\mathcal{H}$ |
|---|---|---|---|---|---|
| Ring based | none | 3 750 994 | 0 | −1876462 | 0 |
| Ring based | BB (GCEPY) | 3 750 592 | 402 | −1876297 | 165 |
| Ring based | F98 | 3 750 570 | 424 | −1876302 | 160 |
| Ring based | Cao13 | 3 750 560 | 434 | −1876276 | 186 |
| Ring based | gNFW$^2$ | 3 750 433 | 561 | −1876232 | 230 |
| Ring based | Coleman20 | 3 750 333 | 661 | −1876144 | 318 |

the ring-based background model with respect to GALPROP-based background models (cf. Tables 2 and 4). Yet the evidence of the GCE, no matter the template adopted, is strong, i.e. $\ln\mathcal{B} \gtrsim 100$. The ranking of GCE models is similar to Table 2, except for the fact of F98 performing worse than Cao13 ($\ln\mathcal{B} = 26$). The gNFW$^2$ template is still preferred over Cao13 ($\ln\mathcal{B} = 44$), while Coleman20 template provides a better fit than the gNFW$^2$ template ($\ln\mathcal{B} = 88$). We, therefore, corroborate the Coleman20 preference found in the GALPROP-based runs, even when the ring-based background model is used.

We notice, however, that the overall goodness of fit of models with the GCE and using the GALPROP$_{8t}$ background template are preferred with respect to our optimization of the ring-based background runs with $\ln\mathcal{B}$ varying between about 500 and 1000 for the different GCE templates. None the less, we will show below (Section 3.4) that these very same models lead to unphysical spectra of the IGRB, questioning their physical interpretations.

Fig. 6 shows the GCE spectra for the five templates tested using both the GALPROP-based and the ring-based background models. Overall, the GCE fluxes are higher when using the gNFW$^2$ template compared with bulge templates, by a factor of a few. All the GCE spectra are relatively soft, and their $E^2 dN/dE$ values peak at around 1–2 GeV.

### 3.4 Model systematics: degeneracy between IGRB and GCE

In Figs 7 and 8, we examine the spectra for the IGRB for the different background models and GCE templates. As can be seen from the top left panel of Fig. 7, when using the GALPROP-based background model, instead, the IGRB is hardly 'detected' around GeV energies, no matter which GCE template is used (including no excess case). This result seems to be unphysical. Comparing the IGRB and GCE flux for different GCE templates (remaining panels of Fig. 7), the IGRB is largely missing around where the GCE is peaked. The issue does not seem to exist for the ring-based background model, which is shown in Fig. 8. The fact that the IGRB in Fig. 7 has no flux around 1–5 GeV is a concerning property of the GALPROP-based background models, which is not seen in the ring-based models. One possible way to ameliorate this issue may be to put strong priors on the IGRB flux so it does not drop to zero.

### 3.5 Analysis systematics: the role of mask size and goodness of fit with Monte Carlo simulations

For the sake of studying the systematic uncertainties induced by the choice of the masked regions, we use the same data and templates as P2022. The main difference with the ring-based case in Section 3.3
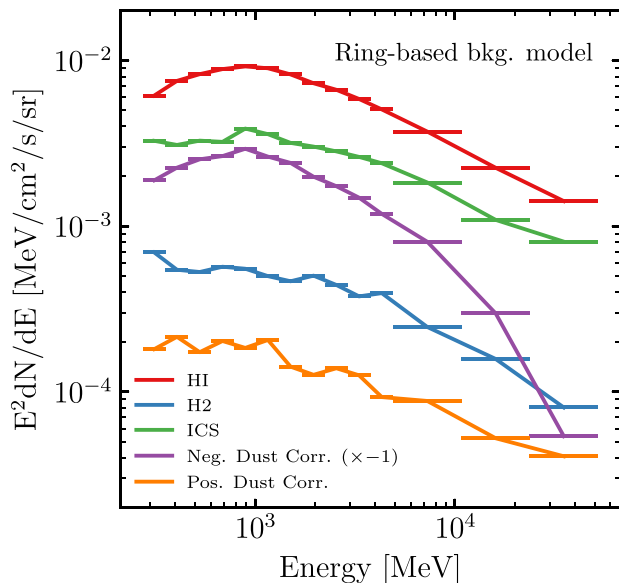
**Figure 4.** Best-fitting spectra for the H I, H2, ICS, negative and positive dust corrections for the ring-based background model (P2022). The results are obtained with no GCE template and are from the L-BFGS-B algorithm.

is that the pixel resolution used is now $0.2°$ rather than $0.1°$. We also use a larger point-source mask, the details of which are shown in Table 5. The new point-source mask was designed to be wide enough to mask out 90 per cent of the flux of each point source in each energy bin. Our masks can be compared to table 1 of the supplementary material of M2023. As can be seen, our point-source masks have larger radii than even their 'large' point-source masks. We found that if we make our masks smaller, then we could see the residual point source signal leaking out from the mask. Note that we do not use the first two energy bins listed in Table 5 in our subsequent analysis for this subsection, as those bins have too high a fraction of the ROI masked out to provide meaningful constraints. In Fig. 9, we compare the fraction of the ROI that is masked in this section to the fraction of the ROI that is masked in earlier sections and in M2023. We also remind the reader that in M2023 and in previous sections, the Galactic plane was masked.

In Table 6, we evaluate the statistical evidence for the additional GCE component, when our new point-source mask, as described above, is applied (but the Galactic plane is unmasked). Since we do not mask the Galactic plane, we also add a model component for the NB (Nishiyama et al. 2013). This table can be compared to table 2 of P2022. As can be seen, the qualitative results are very similar to P2022, with mild difference in the likelihoods probably due to the masking and also to having discarded the first two energy bins so that we have 13 bins rather than the 15 that were used by P2022. As can be seen from the table, we find that, with a larger point-source mask, there is strong evidence for an additional component on top of the ring-based background model. Moreover, there is evidence for the Coleman20 template at $8.1\sigma$ on top of the ring-based + NB model, while the addition of gNFW$^2$ is not significant ($2.8\sigma$). Finally, the significance for DM is strongly reduced to a negligible level when added to the ring-based+NB + Coleman model.

We then run the case with both new point-source mask and Galactic plane mask $|b| < 2°$. In this case, the results are shown in Table 7. As can be seen, with the Galactic plane and new point-source mask, we find neither the gNFW$^2$ DM template nor the NB template

to be significant. Conversely, the Coleman20 BB template is still significant.

P2022 tested the goodness of fit using Monte Carlo simulations. As can be seen from their Fig. 9, the Monte Carlo simulations were not consistent with the fit for the $E < 5$ GeV. This was somewhat ameliorated to $E < 4$ GeV by reducing the ROI from $40°$ by $40°$ to $30°$ by $30°$. In Fig. 10, we show the full ROI Monte Carlo simulations for the new point-source mask, and for the case with both the point sources and Galactic plane masked out respectively. As can be seen, in all energy bins, the Monte Carlo simulations are consistent with the data. This indicates that, with standard template fitting, it is more robust to mask the point sources rather than try to model them when the diffuse Galactic emission is being fit. An alternative would be to include a model of the point sources and simultaneously fit the position of the point sources with the parameters of the Galactic diffuse emission model. However, this would be very computationally intensive and goes beyond the scope of tests necessary in the current analysis.

## 4 ADAPTIVE-TEMPLATE FITTING

In this section, we show that the potential of SKYFACT to reduce residuals and optimize model components in a data-driven way allows for robust inference on the GCE morphology, as it was already shown in the case of the analysis of subthreshold point sources (Calore et al. 2021). Due to a large number of nuisance parameters, it is infeasible to optimize a given gamma-ray emission model on small ROIs, in particular, ROIs with a masked Galactic plane that encompasses the bulk of the detected gamma rays from the GC. Whenever we optimize a model with SKYFACT, we, therefore, perform the optimization with respect to the full ROI of $40° \times 40°$ centred on the GC.

In the SKYFACT-based part of this study, we work with the *Fermi*-LAT data selected according to Table 1 except for a change in bin size from $0.1°$ to $0.25°$ to render the analysis computationally tractable. We consider three distinct model compositions in parallel, namely the GALPROP$_{8t}$, ring-based, and the original SKYFACT (Bartels et al. 2018) backgrounds. In all three gamma-ray emission model set-ups, we employ the Coleman20, NB and gNFW$^2$ templates. Note that in this section, we only consider masking the Galactic plane. The point sources are added to the background model with full spectral freedom per point source. In the following section, we outline how we apply the SKYFACT re-modulation in the context of a masked ROI.

### 4.1 Deriving gamma-ray optimized background models with SKYFACT

SKYFACT enables us to go beyond the model iterations investigated in Section 3 by re-modulating the spatial morphology of the respective model's components (where possible). As stated earlier in Section 2.3.2, this is achieved via adaptive template fitting, invoking a large number of spectral (per energy bin) and spatial modulation parameters (per spatial pixel) whose ranges are controlled by user-input hyperparameters. The degree of variation in these modulation parameters is restricted via a penalizing likelihood function adding to a standard Poisson likelihood term to prevent overfitting.

Such an approach is only feasible with enough information in the considered data set. Thus, applying SKYFACT in the presence of an extensive Galactic plane mask is prohibitive. Therefore, we devise the following scheme to incorporate re-modulated gamma-
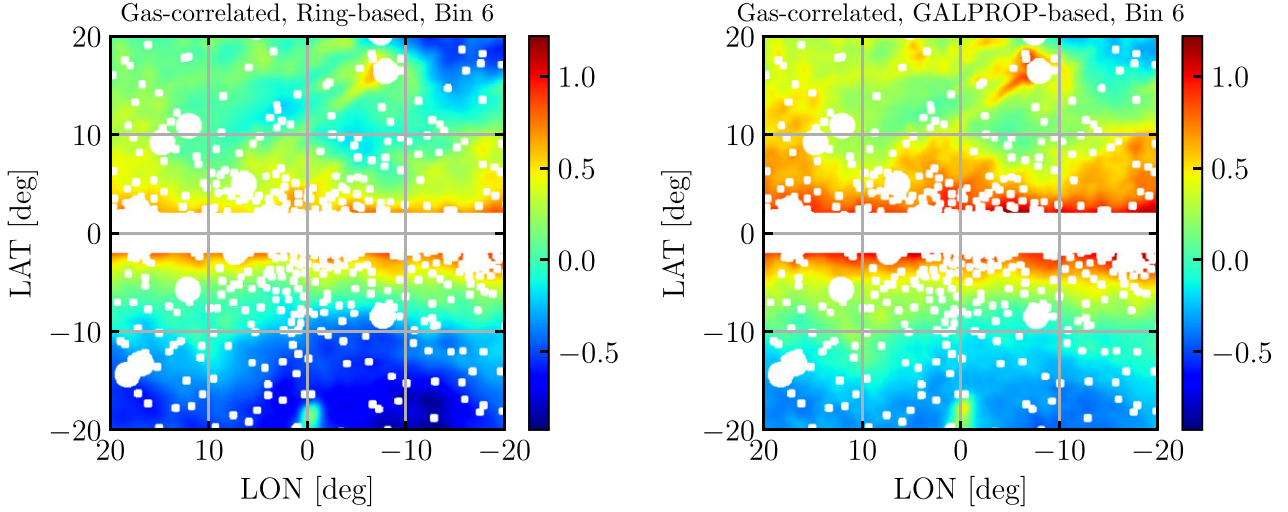
**Figure 5.** Best-fitting count maps (in $\log_{10}$ scale) for the gas-correlated component for the ring-based and GALPROP$_{7p}$ based background model. We show the sixth energy bin (1.02–1.32 GeV).
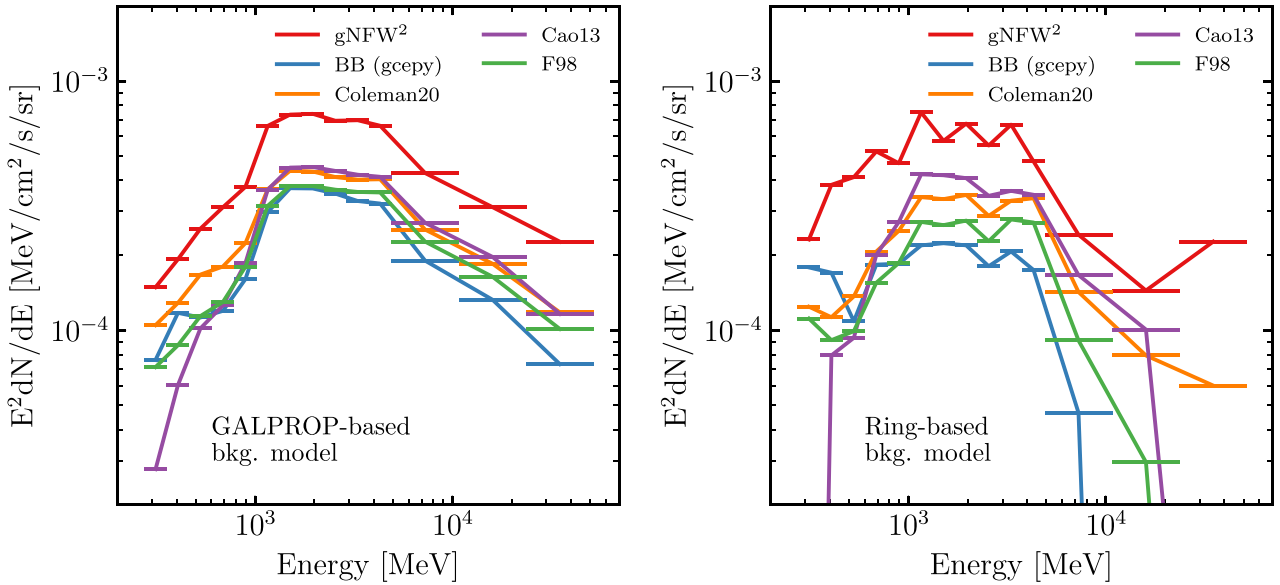


**Figure 6.** GCE spectra for the gNFW$^2$ template as well as the four bulge templates from different references, using the GALPROP$_{8t}$-based background model (left) and the ring-based background model (right).

ray emission models in our analysis. For each of the considered background model set-ups stated above, we perform a fit to the *Fermi*-LAT data – data selection described in Table 1 – by enabling spatial re-modulation. SKYFACT hyperparameter settings are given in Section 2.3.2. We obtain what we call '*optimized*' versions of the original background model set-ups. SKYFACT's optimization will re-modulate the background templates to minimize the residual photons, i.e. parts of the GCE emission will be absorbed by the selection of background components. We deem such an approach conservative because we deliberately diminish the total luminosity of the GCE. However, since the GCE's spatial morphology and spectrum are not fully degenerate with the employed background components, it is very unlikely that the entire excess is re-absorbed

in the optimized background templates (and indeed, we will confirm this with our results). In what follows, we investigate the performance of these optimized background model iterations on data sets with a Galactic plane mask utilizing, in a second step, standard template fits.

To convey an idea of how the SKYFACT optimized models compare to the original versions, we explicitly go through the derivation of the optimized version of P2022. In Table 8, we list the astrophysical gamma-ray emission model components of the original set-up of P2022, which we employ in this part of the analysis. The selected spatial templates and associated spectra guarantee the optimization of each component based on physical priors. The GCE components, in particular, are initialized following the average spectrum of Galactic
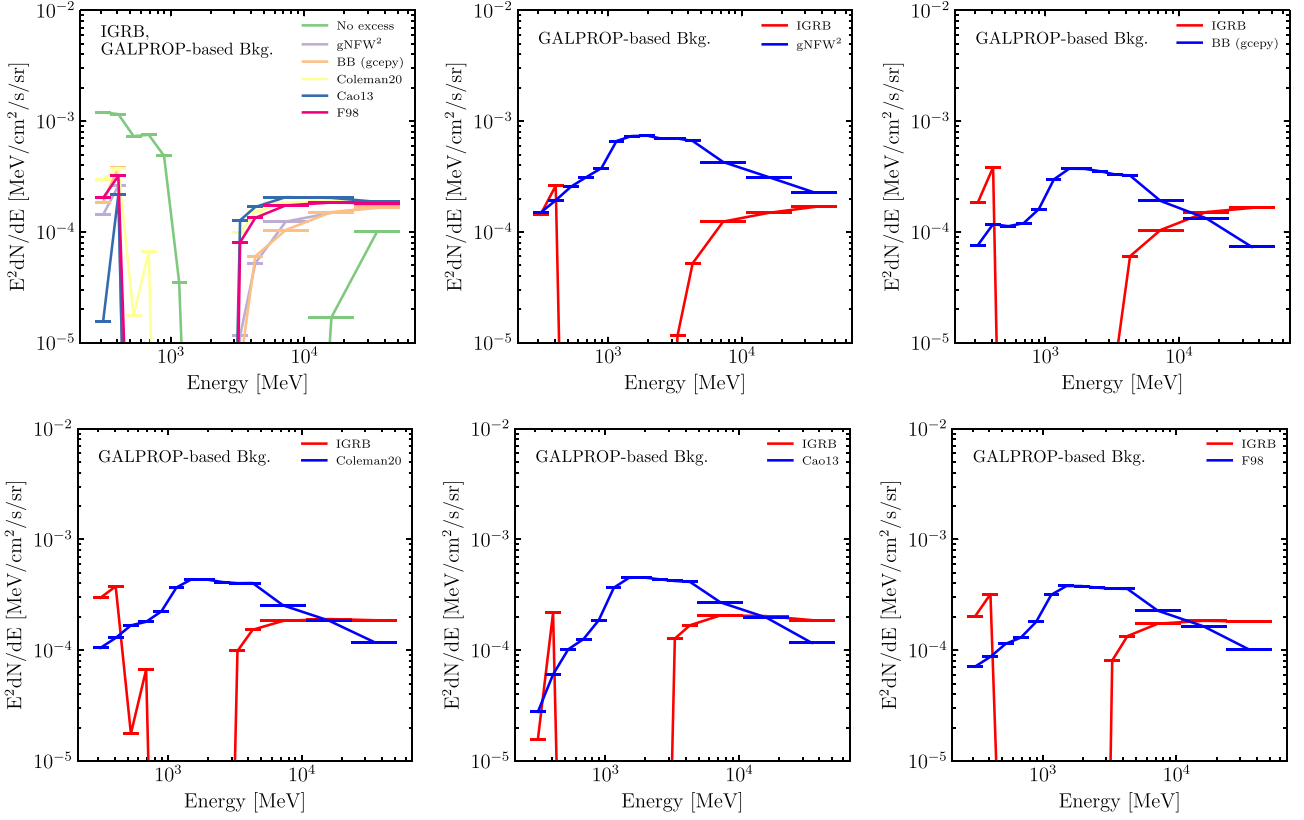
**Figure 7.** Top left: Spectra of IGRB in the GALPROP-based background model, using different GCE templates (including no excess). *Remaining*: Comparing the IGRB and GCE fluxes when different GCE templates are used in the GALPROP-based background model.

MSPs detected by *Fermi*-LAT (McCann 2015), which suppresses the gamma-ray emission above a few tens of GeV.[7]

To reduce the required computation time of SKYFACT and to improve its convergence, we combine the H I and H2 templates per ring into a single component, and we create a single inverse Compton template from the six initial rings. The results of the optimization run – in this case using the example of modelling the GCE with Coleman20, NB, and gNFW$^2$ to exemplify what is maximally possible regarding the reduction of fit residuals – are shown in the right panel of Fig. 11 in terms of significance ((data − model)/$\sqrt{\text{model}}$) of the residuals in the energy range from 1.72 to 10.8 GeV. We compare these residuals to the residuals we obtain with a simple template fit (the left panel of the same figure) using the full model as described in Table 8, and which should correspond to results in Section 3.3. Here, a template fit refers to a maximum likelihood fit where the spatial morphology of all templates is fixed, i.e. turning off all spatial and spectral re-modulation parameters, while all spectra are completely unconstrained and free to vary. As intended, SKYFACT is able to noticeably reduce the significant residuals of the template fit along the Galactic plane to the left and right of the GC. Moreover, the remaining residuals of the optimized model appear rather featureless and well-distributed around zero. We obtain very similar results when the GCE is modelled with only the gNFW$^2$ DM template. This optimization procedure is repeated for the remaining background models, M2023's GALPROP$_{8t}$ and the model setup of RUN5 used in the original SKYFACT works.

---

[7]Further exploration of the GCE properties above 10 GeV with SKYFACT is presented in Manconi, Calore & Donato (2024).

## 4.2 DM evidence in masked analyses with the original SKYFACT set-up.

To investigate the preferred spatial morphology of the GCE, we turn towards an alternative gamma-ray emission model neither probed in P2022 nor M2023. The RUN5 model iteration compiled for the original SKYFACT works (Storm et al. 2017; Bartels et al. 2018) provides an ideal candidate to shed light on the impact of a Galactic plane mask and its impact on template-based fits. To this end, we fully adopt the set-up of run5 in terms of model composition (see the cited publications for all details) and SKYFACT's hyperparameter settings. This gamma-ray emission model iteration contains representatives of most of the components listed in Table 8 except for the dust correction, Loop I, Sun, and Moon contributions.

In contrast to the other two background model iterations, this one can only be used in its 'optimized' version as the original templates of the model are not meant to fit the data well. Following our rationale outlined in the previous section, we first perform an optimization run with respect to the selected *Fermi*-LAT dataset without masking any part of the sky while adding no explicit GCE components to the model definition. Afterwards, we extract the optimized model components with the aim of conducting several template fits based on the optimized background templates with varying Galactic plane mask sizes. In these template fits, we restrict the full energy range of the selected *Fermi*-LAT data set to energy bins covering 500 MeV to 12 GeV, i.e. 10 energy bins in total, in order to save computation time while still capturing the bulk of the GCE's emission.

To derive a statistically sound assessment of the data's preference for any particular GCE morphology, we need to perform template fits of nested models. We perform this type of model comparison in
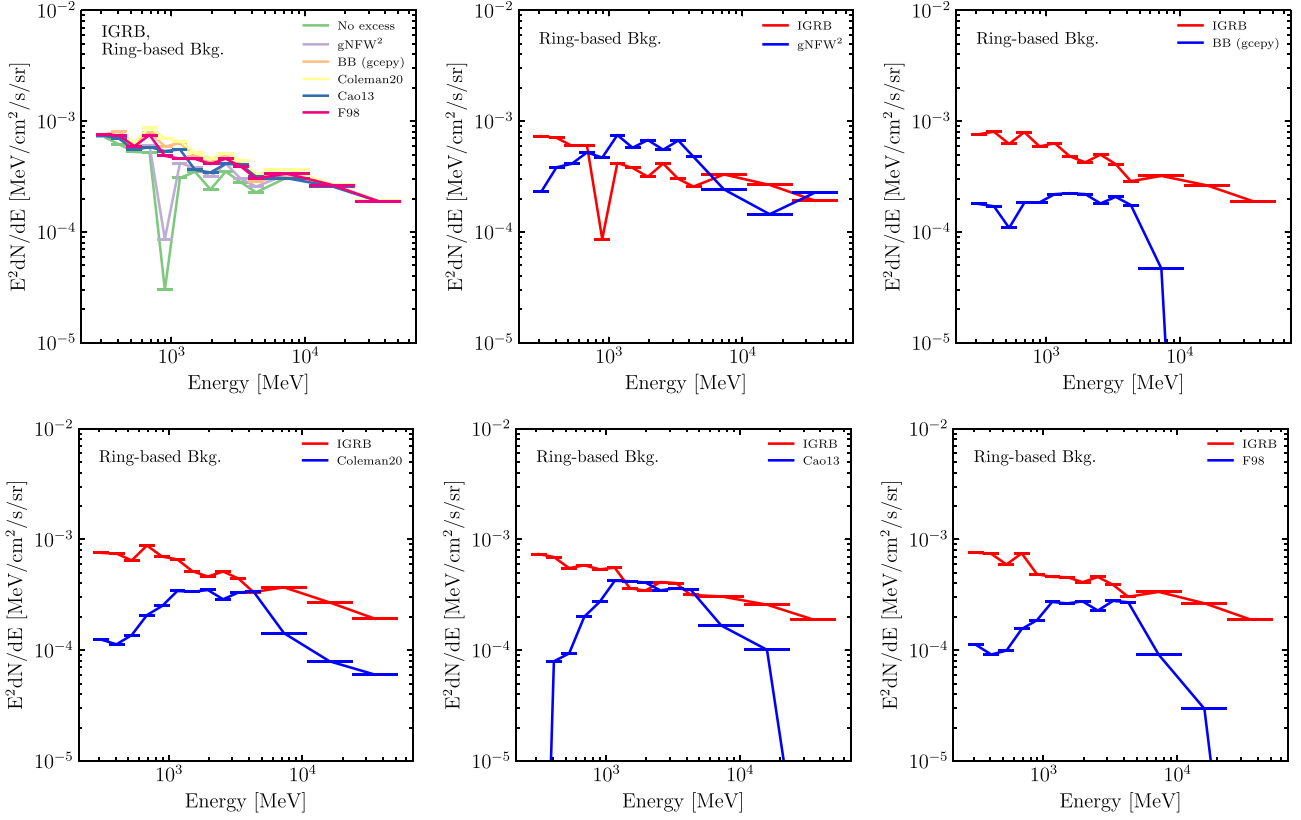
**Figure 8.** Same as Fig. 7, but for the ring-based background model.

**Table 5.** The energy bins and the radii, $\theta$, of the point source masks used for Section 3.5. The last column shows the fraction of pixels masked relative to the total number of pixels in the inner $40° \times 40°$ GC region, including both the 4FGL-DR2 catalogue masked point sources and the Galactic plane $|b| < 2°$ mask.

| $E_{min}$–$E_{max}$ (GeV) | $\theta$ (°) | Masked fraction (per cent) |
|---|---|---|
| 0.667–0.889 | 1.92 | 88.5 |
| 0.889–1.19 | 1.58 | 80.3 |
| 1.19–1.58 | 1.28 | 68.7 |
| 1.58–2.11 | 1.04 | 58.9 |
| 2.11–2.81 | 0.8 | 49.0 |
| 2.81–3.75 | 0.72 | 41.0 |
| 3.75–5.0 | 0.56 | 35.3 |
| 5.0–6.67 | 0.48 | 28.5 |
| 6.67–8.89 | 0.36 | 26.1 |
| 8.89–11.9 | 0.32 | 20.7 |
| 11.9–15.8 | 0.2 | 17.7 |
| 15.8–21.1 | 0.2 | 17.7 |
| 21.1–28.1 | 0.2 | 17.7 |
| 28.1–37.5 | 0.2 | 17.7 |
| 37.5–158 | 0.2 | 17.7 |



**Figure 9.** A comparison of the fraction of the sky masked for the analysis in of the sky masked in Section 3.5 (blue) and the fraction of sky masked for the rest of Section 3 and also M2023 (red). Note that in both cases, the plot is for the combined point source and Galactic plane mask.

terms of the significance of the additional component as explained in Section 2.3.3. We define our base model as all astrophysical background components plus a GCE represented by Coleman20 and NB templates. An extended model adds the gNFW² template so that we can compare the likelihood values for fits with both model
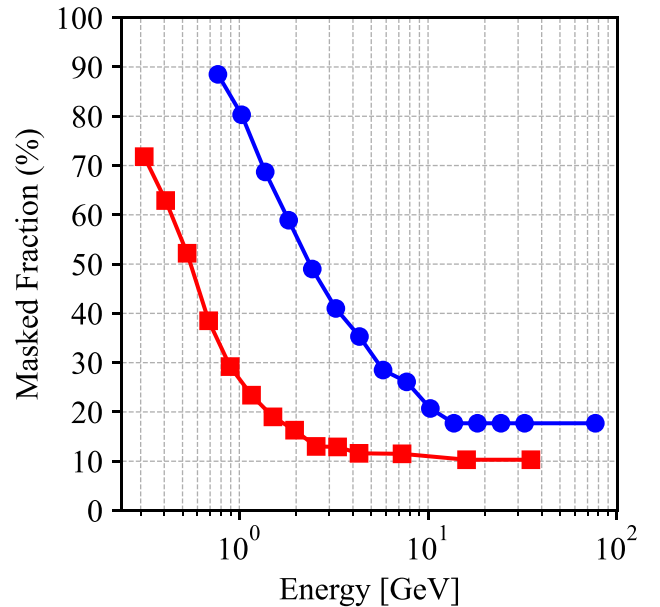
**Table 6.** Statistical significance of the GCE templates for the ring-based background model of P2022 when the new point-source mask is applied (plane unmasked). Additional sources considered in the analysis are NB, Coleman20 BB, and gNFW$^2$ DM-like template. In addition to the TS, the significance of the additional component is also given in terms of the equivalent number of $\sigma$.

| Baseline model | Additional source | TS | Significance |
|---|---|---|---|
| Ring based | Coleman20 | 77.5 | 7.3 $\sigma$ |
| Ring based | gNFW$^2$ | 80.7 | 7.5 $\sigma$ |
| Ring based | NB | 299.7 | 16.2 $\sigma$ |
| Ring based + NB | gNFW$^2$ | 21.0 | 2.8 $\sigma$ |
| Ring based + NB | Coleman20 | 90.9 | 8.1 $\sigma$ |
| Ring based+NB + Coleman20 | gNFW$^2$ | 3.5 | 0.3 $\sigma$ |

**Table 7.** The same as Table 6 except that the Galactic plane ($|b| < 2°$) is also masked.

| Baseline model | Additional source | TS | Significance |
|---|---|---|---|
| Ring based | gNFW$^2$ | 12 | 1.7 $\sigma$ |
| Ring based | NB | 19 | 2.6 $\sigma$ |
| Ring based | Coleman20 | 56 | 5.9 $\sigma$ |
| Ring based + Coleman20 | NB | 3 | 0.2 $\sigma$ |
| Ring based + Coleman20 | gNFW$^2$ | 5 | 0.5 $\sigma$ |

instances. Adding a gNFW$^2$ component to the base model essentially adds 11 degrees of freedom or parameters, which must be strictly positive, namely the normalizations of the DM template per energy bin and a global normalization parameter.

The results of this approach are reported in Table 9. The first row of this table can be compared to the last row of table 2 of P2022. There a slightly lower significance was found but some minor differences are to be expected given variations on how the point sources are treated, the number of energy bins used, and the modulation done of the background templates. In both cases, the significance of a gNFW$^2$ template is negligible once the Coleman20 BB and NB have been added. The results of Table 9 indicate that the evidence for the necessity of adding a gNFW$^2$ template to the gamma-ray emission model is at most 1.6$\sigma$ in the case of no Galactic plane mask. Interpreted differently, there is only marginal evidence that the preferred morphology of the GCE follows a gNFW$^2$ profile. This is in agreement with results of Section 3.5, also when spatial modulation of the background model components is allowed.

## 4.3 Bayesian model comparison of original and optimized gamma-ray emission models

Given the sometimes opposing findings on the preferred spatial morphology of the GCE reported in the broad body of literature, it is necessary to ask the question of how much the employed astrophysical background model impacts the eventual conclusion. Here, we investigate this question from a Bayesian point of view by quantifying the degree of belief in certain gamma-ray emission models, that is, what model fits the *Fermi*-LAT data best. The expectation is to verify that with increasing Bayesian evidence for a gamma-ray emission model, the preference for a particular spatial morphology of the GCE is converging to either DM represented by a gNFW$^2$ template or the combination of the Coleman20 BB and NB templates.

### 4.3.1 Model comparison without GCE components.

In Table 10, we consider six background models, three original background model template sets and three optimized template sets obtained by applying the rationale outlined in Section 4.1. We first compare the performance of these models in a template fit without adding additional GCE components, i.e. we assess how well the background templates alone fit the GC gamma-ray emission. Note that we also include M2023's GALPROP$_{7p}$ background model (only original templates) since it was found in M2023 that it yields the best fit when not accounting for GCE components. In all runs, we applied a Galactic plane mask of $|b| < 2°$, which turned out to be a crucial ingredient in the comparisons of DM-like and bulge templates.

To derive the stated Bayesian evidence $\mathcal{H}$, we proceed as follows. We extract the best-fitting template normalizations for each model iteration. We sum all model components multiplied by the retrieved best-fitting normalizations. Using MULTINEST (Feroz, Hobson & Bridges 2009) and specifying 1000 live points and an evidence tolerance of 0.2, we re-fit the masked ROI with these models while assigning a single normalization parameter to it and employing a Poisson log-likelihood function.

By comparing these models,[8] we notice that the optimized versions of each respective background model iteration always yield a better fit to the data, also in the case of ring-based background models. In contrast to M2023, we do not find that the original GALPROP$_{7p}$ performs better than the original GALPROP$_{8t}$ without additional GCE components as shown by a Bayes factor of $\ln \mathcal{B} = 277$ in favour of the latter.

Among all considered model iterations, we find that the run5 model of the original SKYFACT works yields the best description of the data by far ($\ln \mathcal{B} > 1000$ to the next best iteration, the optimized M2023's GALPROP$_{8t}$). Yet, already at this stage, we caution the reader not to overinterpret the quoted evidence values $-\ln \mathcal{H}$. skyFACT is not a perfect tool and cannot re-modulate the templates to 100 per cent accuracy. For example, when we compare the optimized templates associated with the $\pi^0$ and bremsstrahlung emission (following the gas distribution in the Milky Way) among different background model iterations, we find that they do not converge to the exact same morphology. For instance in the most extreme case, the relative deviation of this optimized gas-related component is on average around 30 per cent with respect to SKYFACT's RUN5 and M2023's GALPROP$_{8t}$. On one side, this is caused by the quite diverse gamma-ray emission model composition in P2022, M2023 and SKYFACT's RUN5, which yields different priors for the adaptive template fitting routine. On the other side, this technique relies on user-defined hyperparameters that alter the final results. While we improve the fit results via SKYFACT, we do not claim to have derived the unique optimal diffuse model describing the physics of the GC.

---

[8]We stress that the numbers presented in Table 10 cannot and should not be directly compared to the previous results in Section 3. The main reason is the differing resolution of the templates adopted for the two analyses. In Section 3, we use a finer resolution of 0.1°, which directly translates to larger likelihood values. At the same time, we have to project the gamma-ray flux models of each component to the chosen geometry of the data set. The initial resolution of the flux model is most of the times coarser than the chosen bin size for the data. Projection effects can distort and washout information so that we do not compare the exact same model each time we change the resolution.
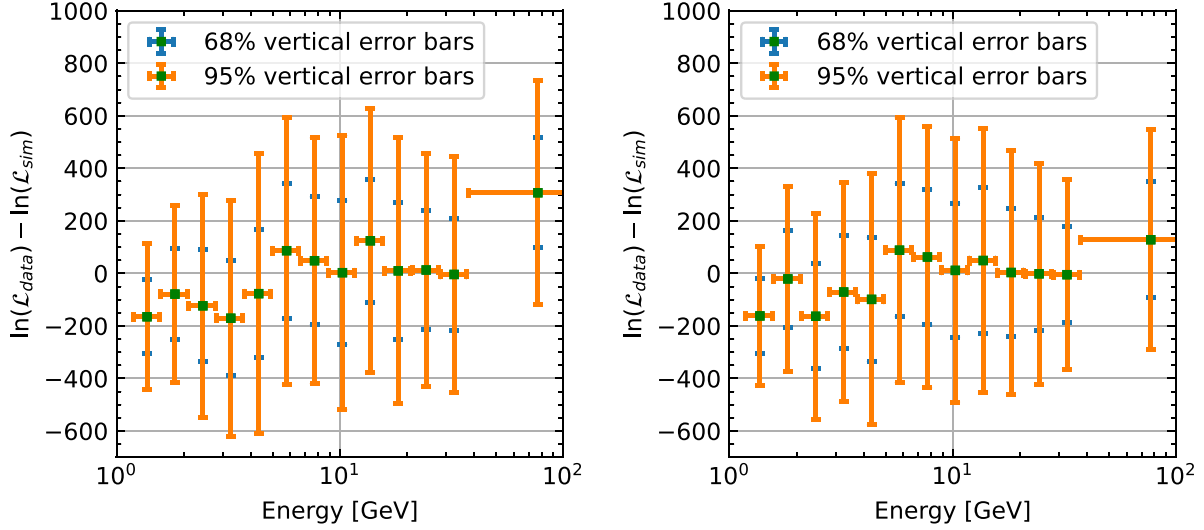
**Figure 10.** Left: Monte Carlo simulations for the ring-based+NB + Coleman20 model when the point sources are masked as specified in Table 5. Poissonian simulations were generated for the best-fitting parameters of this model. Each simulation was fit using the ring-based+NB + Coleman20 model, and the maximum likelihood for the simulation ($\mathcal{L}_{sim}$) was compared to the maximum likelihood for the ring-based+NB + Coleman20 model fit to the *Fermi*-LAT data (ln $\mathcal{L}_{data}$) which is given in Table 6. The vertical error bars were estimated from the mean and standard deviation of the simulation samples. The horizontal error bars indicate the energy bin widths. Right: Same as the left plot except that the Monte Carlo simulations are for the ring-based + Coleman20 model when the point sources are masked as specified in Table 5 and the $|b| < 2°$ Galactic plane was also masked.

**Table 8.** Summary of the spatial and spectral model components used as input for SKYFACT to derive an optimized version of the gamma-ray emission model of P2022.

| Component | Spatial morphology | Spectrum |
|---|---|---|
| $\pi^0$ + bremsstrahlung | H I maps P2022 and H2 maps Macias et al. (2018) in four rings (in Galactocentric radii: 0–3.5, 3.5–8, 8–10, 10–50 kpc) | Ackermann et al. (2012) |
| Inverse Compton | Numerical computation in six Galactocentric rings via GALPROP v56 | Spectrum of foreground |
| | Based on the 3D interstellar radiation field model of (Porter, Johannesson & Moskalenko 2017) | Model A (Ackermann et al. 2015) |
| Dust correction | Positive and negative corrections maps of Abdollahi et al. (2020) | Power law $\propto E^{-2}$ |
| Detected sources | 4FGL-DR2 sources in our ROI (Ballet et al. 2020) | Spectra listed in 4FGL-DR2 |
| Isotropic gamma-ray background | Isotropic | Fermi Science Tools |
| Fermi bubbles | Macias et al. (2019) | Ackermann et al. (2017) for low latitudes |
| Sun and Moon | Data-driven, derived with the Fermi Science Tools | Fermi science tools |
| Loop I | Wolleben (2007) | Power law $\propto E^{-2}$ |
| DM | C2022 | $\propto (E/1\,\mathrm{GeV})^{-1.46} \exp(-E/3.6\,\mathrm{GeV})$ |
| Boxy Bulge | Coleman et al. (2020) | $\propto (E/1\,\mathrm{GeV})^{-1.46} \exp(-E/3.6\,\mathrm{GeV})$ |
| Nuclear stellar cluster | Nishiyama et al. (2013) | $\propto (E/1\,\mathrm{GeV})^{-1.46} \exp(-E/3.6\,\mathrm{GeV})$ |

### 4.3.2 Model comparison with added stellar GCE components

For each of the all gamma-ray emission models in Table 10 but GALPROP$_{7p}$ original, we then add an additional GCE component, modelled as Coleman20 and NB, perform a standard template fit and extract Bayesian evidence values as described in the previous paragraph.

As can be seen from Table 11, in general, we find very strong evidence for the combination of the Coleman20 BB and NB on top of the background-only model iterations, as clearly implied by the large significance values around $20\sigma$ in all tested cases. This means that, regardless of the SKYFACT optimization, *Fermi*-LAT data strongly want an additional component, i.e. the GCE. As can be seen from Table 11, among the 'original' non-SKYFACT modulated gamma-ray emission models, we find that the one proposed by M2023, GALPROP$_{8t}$, exhibits strong evidence (ln $B \approx 300$) for being a better fit to the data with respect to the original P2022 set-up. This claim

has been made in M2023, which we are able to reproduce here (and in Section 3). However, this does not mean that their model is, in general, the best description of reality since we are only looking at latitudes $|b| > 2°$. The SKYFACT-modulated versions of both gamma-ray emission model instances are strongly preferred by the data compared to their original counterparts. Globally, the SKYFACTRUN5 model is the one performing best among all tested cases, even when spatial modulation are allowed on the original M2023 and P2022 models. There is strong evidence for it being preferred over the second-best model, M2023 optimized, by ln $\mathcal{B} \approx 1400$.

### 4.3.3 Model comparison and significance of a DM component

Finally, we added a gNFW$^2$ component in the subsequent fit, on top of the bulge Coleman20 and NB model. This way, we can repeat the approach outlined in Section 2.3 to quantify the statistical
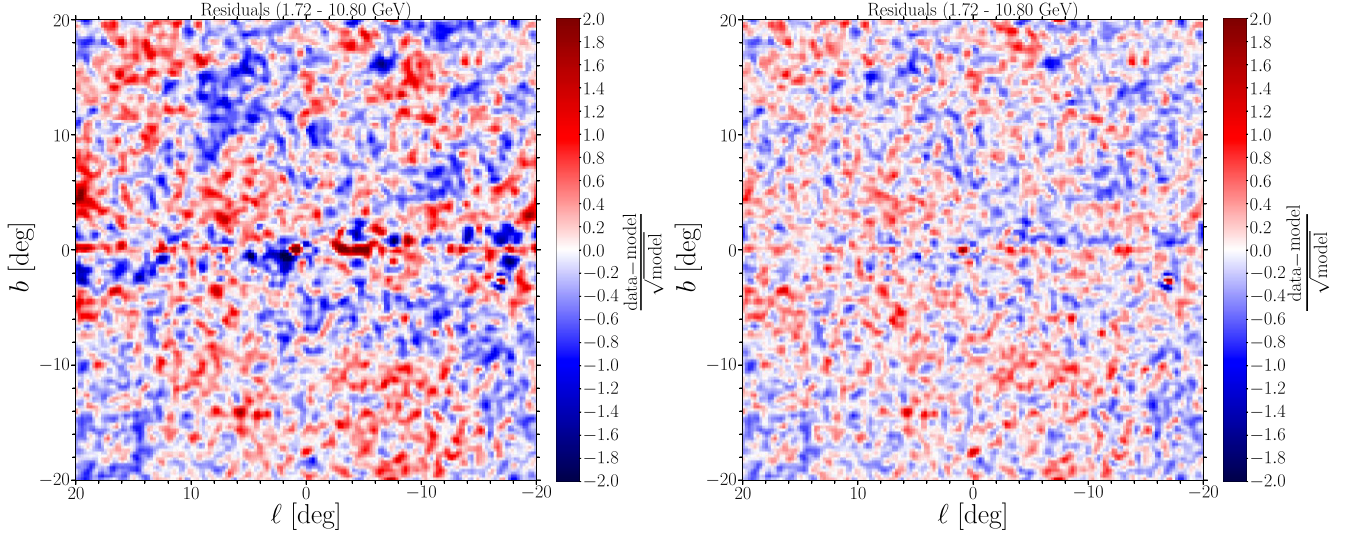
**Figure 11.** Shows (data − model)/√model of residuals in the integrated energy range from 1.72 to 10.8 GeV. Left: Employing SKYFACT to perform a template fit with the ring-based astrophysical model components listed in Table 8 with a GCE represented by the Coleman20 + NB + gNFW². Right: Running the same model set-up with the full re-modulation power of SKYFACT to optimize the employed components.

**Table 9.** Summary of the significance of an additional gNFW² template in template fits with the SKYFACTRUN5 set-up for various Galactic plane mask sizes. Here, 'base' refers to the SKYFACT optimized RUN5 background templates plus the Coleman20 BB and a NB component. The SKYFACT optimization of background templates was performed on the RUN5 set-up without GCE components as outlined in Section 4.1.

| Galactic plane mask | $-2\ln\mathcal{L}_{\text{base+gNFW}^2}$ | $-2\ln\mathcal{L}_{\text{base}}$ | TS | Significance |
|---|---|---|---|---|
| no mask | 275 081 | 275 071 | 10 | $1.6\sigma$ |
| $|b| < 1°$ | 260 989 | 260 985 | 4 | $0.5\sigma$ |
| $|b| < 2°$ | 247 030 | 247 029 | 1 | $0.1\sigma$ |
| $|b| < 5°$ | 205 937 | 205 932 | 5 | $0.7\sigma$ |

**Table 10.** Summary of the likelihoods ($\mathcal{L}$) and evidence ($\mathcal{H}$) for different background models. A Galactic plane mask of $|b| < 2°$ is applied, but the point sources are included as part of the background model. The optimized version of the background model includes a skyFACT modulated version of the non-point source components of the 'original' model.

| Background model | $-2\ln(\mathcal{L})$ | $-\ln(\mathcal{H})$ |
|---|---|---|
| M2023's GALPROP₇ₚ 'original' | 347 477 | 280175 |
| M2023's GALPROP₈ₜ 'original' | 347 465 | 279898 |
| M2023's GALPROP₈ₜ 'optimized' | 342 008 | 274036 |
| P2022's ring-based 'original' | 346 859 | 279723 |
| P2022's ring-based 'optimized' | 342 982 | 276 075 |
| SKYFACTRUN5 'optimized' | 340 266 | 272900 |

significance of an additional gNFW² template from the frequentist perspective. As can be seen from Table 12, we are also able to reproduce the M2023 result of the strong evidence (more than $11\sigma$) for the necessity of an additional DM gNFW² template on top of Coleman20 and NB in the context of the original M2023 model set-up. However, as can also be seen from Table 12, the significance of the gNFW² component is only marginal after SKYFACT modulating the spatial morphology of the background model. In contrast, consistent with earlier works and the results in Table 7, we see in Table 12 that the setup of P2022 never required an additional gNFW² template

after accounting for the GCE as the Coleman20 BB template and NB template. As can also be seen from Table 12, the SKYFACT model is an outlier here because the fit, including a gNFW² template, is even worse than the one without. This situation can occur in SKYFACT since even in a template fit, we modulate the spatial morphology of the detected extended 4FGL-DR2 sources. Thus, the penalizing likelihood function adds a non-vanishing part to the overall value of the likelihood function. Yet, the extensions of these sources are marginal compared to the rest of the ROI, so we do not expect biased results. Consequently, there is also no evidence of the need for a gNFW² template in this gamma-ray emission model iteration.

In conclusion, whenever we employ an optimized astrophysical background model, there is no strong evidence for spherical symmetry of the GCE or at least a preference for such a morphology even when masking the Galactic plane. Previous contrary findings seem to be driven by a certain amount of background mismodelling. We stress again that quoted values of the Bayesian evidence are subject to the caveats raised in Section 2.3.3. Eventually, the Bayesian framework allowed us to single out the SKYFACTRUN5 set-up to be the most suitable to describe the Galactic centre physics with adaptive template fitting. Its assumed priors for the spatial and spectral profile of the used components yield the best-fitting model among the tested cases, although it does by no means imply that it is the optimal model achievable.

## 5 CONCLUSIONS

We have performed an extensive analysis of models of gamma-ray emission towards the GC as an explanation of the GCE, subject to different choices of diffuse background models, point source and Galactic plane masking, and extended source models. In particular, we tested GALPROP-based background models versus more flexible non-parametric ring-based models.

First, we have thoroughly tested contradicting results in the literature for masked analyses, especially those pertaining to the preference for stellar bulge versus DM, e.g. in M2023, where preference for a gNFW² (DM-like) emission of the GCE was found.

**Table 11.** Summary of the TSs, significances for the Coleman20 + NB templates for different background models. The evidence ($\mathcal{H}$) is also given for the combined model of background + Coleman20 + NB. See Table 10 for an explanation of the background models and their corresponding likelihoods.

| Background model | TS (Coleman20 + NB) | Significance | $-\ln(\mathcal{H})$ |
|---|---|---|---|
| M2023's GALPROP$_{8t}$ 'original' | 2137 | $> 20\sigma$ | 278 929 |
| M2023's GALPROP$_{8t}$ 'optimized' | 4822 | $> 20\sigma$ | 272 018 |
| P2022's ring-based 'original' | 445 | $19.1\sigma$ | 279 231 |
| P2022's ring-based 'optimized' | 5435 | $> 20\sigma$ | 272 573 |
| SKYFACTRUN5 'optimized' | 3957 | $> 20\sigma$ | 270 627 |

**Table 12.** Summary of the TSs and significances for the gNFW$^2$ template for different background models once the Coleman20 + NB templates have been added. See Table 10 for an explanation of the background models and their corresponding likelihoods.

| Base model | TS(gNFW$^2$) | Significance | $-\ln(\mathcal{H})$ |
|---|---|---|---|
| M2023's GALPROP$_{8t}$ 'original' + Coleman20 + NB | 162 | $11.4\sigma$ | 278 843 |
| M2023's GALPROP$_{8t}$ 'optimized' + Coleman20 + NB | 22 | $2.7\sigma$ | 272 008 |
| P2022's ring-based 'original' + Coleman20 + NB | 32 | $3.9\sigma$ | 279 202 |
| P2022's ring-based 'optimized' + Coleman20 + NB | 14 | $1.7\sigma$ | 272 504 |
| SKYFACTRUN5 'optimized' + Coleman20 + NB | $-3$ | – | 270 673 |

In Section 3, we showed that we can reproduce the analysis in M2023, and we scrutinized their main results in light of model and analysis systematic uncertainties. In Section 3.2, we highlighted the relevance of the bulge templates: when using the same GALPROP-based background models as in M2023, with the same data selection and masks, we demonstrated using Bayesian evidence that the Coleman20 and F98 bulge models provide a better description of the inner Galaxy gamma-ray sky than a gNFW$^2$ model. We then tested an alternative model for the Galactic diffuse emission, and, in particular, the so-called ring-based background model. In Section 3.3, we demonstrated that, in the absence of an additional GCE source, the ring-based model better performs with respect to GALPROP-based models. When adding a GCE source, the Coleman20 bulge model is the preferred model of the GCE, significantly better than the gNFW$^2$ template. We notice that the contrary conclusions of M2023 when using the ring-based model were due to their not correctly finding the minimum of the parameters for the ring-based templates due to an overly restrictive prior. We also found that M2023 used a non-standard version of the Galactic bulge template. This is confirmed by the fact that even with their GALPROP-based templates, the better-motivated Coleman20 bulge template still provides a superior fit to the *Fermi*-LAT data. In Section 3.5, we examined the case of ring-based templates with more aggressive masking. We found that, when just the point sources are masked out, the Coleman20 BB and NB significantly improve the fit, and once they are added, the gNFW$^2$ template does not significantly improve the fit anymore. When the Galactic plane is masked out, only the Coleman20 BB template is required. The NB is indeed too small in its spatial extent to have any significant effect on the model fit to the data in that case. We also showed using Monte Carlo simulations that the fits were consistent with simulations.

We then looked at the case where templates could be spatially modulated using SKYFACT. Allowing for more freedom on the spatial parts of the model components, we were able to further minimize the residuals and improve the goodness of fit. We compared the different background models from the previous sections and added an additional SKYFACT model (RUN5). We demonstrated that switching on the spatial modulation of background models always provide a better fit to data, because of reduction of the residuals. Among all considered background model iterations, we find that the RUN5 model

of the original SKYFACT works yielded the best description of the data by far ($\ln \mathcal{B} > 1000$ to the next best iteration, the optimized M2023's GALPROP$_{8t}$). None the less, limitations still exist in the current SKYFACT implementation. While we improve the fit results via SKYFACT, we do not claim to have derived the unique optimal diffuse model describing the physics of the GC. No matter what the background model is, we always found strong evidence for the Coleman20 + NB model. Moreover, for all background optimized models there is no additional evidence for a DM-like signal. We found similar results, i.e. DM evidence on top of the bulge model always below the $4\sigma$ threshold, for almost all background models. We encountered one exception, namely the original M2023 GALPROP-based template. However, this was found to have a much lower Bayesian evidence in comparison to the SKYFACTRUN5 model. Finally, we found that, for the RUN5 model of the original SKYFACT implementation, the evidence for an additional DM-like contribution is not significant on top of the Coleman20 bulge and NB model regardless of the cut on Galactic latitude.

We stress that throughout this work we have adopted Bayesian statistics when performing model comparison, and our conclusions have to be interpreted in such a statistical framework.

In summary, the preference for a bulge-like morphology of the GCE in the various analyses we have done puts on even more solid grounds the possibility that part of the excess originates from unresolved point sources, such as MSPs. Future multiwavelength analyses of the GC will help determine the nature of the sources emitting across the multi-messenger spectrum from radio (Calore et al. 2016), X-rays (Berteaud et al. 2021), up to very high energy gamma rays (Song, Macias & Horiuchi 2019; Macias et al. 2021).

Note: While our article was near completion, a new article came out (Zhong & Cholis 2024) which found that when the GALPROP-based background model was used, the Coleman20 bulge had a similar likelihood to the gNFW$^2$. No Bayesian model comparison is performed therein.

on the manuscript. The work of SH is supported by the US Department of Energy Office of Science under award number DE-SC0020262, NSF Grant numbers AST1908960 and PHY-2209420, and JSPS KAKENHI Grant number JP22K03630 and JP23H04899. The work of MK and KA is supported by NSF grant no. PHY-2210283. This work was supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. FC and CE acknowledge support by the 'Agence Nationale de la Recherche', grant no. ANR-19-CE31-0005-01 (PI: F. Calore). The work of CE has been supported by the EOSC Future project which is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, grant agreement 101017536. CE further acknowledges support from the COFUND action of Horizon Europe's Marie Skłodowska-Curie Actions research programme, grant agreement 101081355 (SMASH). The work of DS is supported by JSPS KAKENHI grant number 20H05852.

## DATA AVAILABILITY

The templates and some of the software used in this analysis may be obtained by contacting the authors.

## REFERENCES

Abazajian K. N., 2011, J. Cosmol. Astropart. Phys., 2011, 010
Abazajian K. N., Kaplinghat M., 2012, Phys. Rev. D, 86, 083511
Abazajian K. N., Canac N., Horiuchi S., Kaplinghat M., 2014, Phys. Rev. D, 90, 023526
Abazajian K. N., Horiuchi S., Kaplinghat M., Keeley R. E., Macias O., 2020, Phys. Rev. D, 102, 043012
Abdollahi S. et al., 2020, ApJS, 247, 33
Acero F. et al., 2016, ApJ, 223, 26
Ackermann M. et al., 2012, ApJ, 750, 3
Ackermann M. et al., 2015, ApJ, 799, 86
Ackermann M. et al., 2017, ApJ, 840, 43
Ajello M. et al., 2016, ApJ, 819, 44
Atwood W. B. et al., 2009, ApJ, 697, 1071
Ballet J., Burnett T. H., Digel S. W., Lott B., 2020, preprint (arXiv:2005.11208)
Bartels R., Krishnamurthy S., Weniger C., 2016, Phys. Rev. Lett., 116, 051102
Bartels R., Storm E., Weniger C., Calore F., 2018, Nat. Astron., 2, 819
Berteaud J., Calore F., Clavel M., Serpico P. D., Dubus G., Petrucci P.-O., 2021, Phys. Rev. D, 104, 043007
Binney J., Gerhard O. E., Stark A. A., Bally J., Uchida K. I., 1991, MNRAS, 252, 210
Bland-Hawthorn J., Gerhard O., 2016, Ann. Rev. A and A, 54, 529
Buschmann M., Rodd N. L., Safdi B. R., Chang L. J., Mishra-Sharma S., Lisanti M., Macias O., 2020, Phys. Rev. D, 102, 023023
Calore F., Cholis I., Weniger C., 2015, J. Cosmol. Astropart. Phys., 2015, 038
Calore F., Di Mauro M., Donato F., Hessels J. W. T., Weniger C., 2016, ApJ, 827, 143
Calore F., Donato F., Manconi S., 2021, Phys. Rev. Lett., 127, 161102
Cao L., Mao S., Nataf D., Rattenbury N. J., Gould A., 2013, MNRAS, 434, 595 (Cao13)
Caron S., Eckner C., Hendriks L., Jóhannesson G., Ruiz de Austri R., Zaharijas G., 2023, J. Cosmol. Astropart. Phys., 2023, 013
Chang L. J., Mishra-Sharma S., Lisanti M., Buschmann M., Rodd N. L., Safdi B. R., 2020, Phys. Rev. D, 101, 023014
Cholis I., Zhong Y.-M., McDermott S. D., Surdutovich J. P., 2022a, The Return of the Templates: Revisiting the Galactic Center Excess with Multi-Messenger Observations, https://doi.org/10.5281/zenodo.6423495
Cholis I., Zhong Y.-M., McDermott S. D., Surdutovich J. P., 2022b, Phys. Rev. D, 105, 103023 (C2022)
Coleman B., Paterson D., Gordon C., Macias O., Ploeg H., 2020, MNRAS, 495, 3350(Coleman20)
Daylan T., Finkbeiner D. P., Hooper D., Linden T., Portillo S. K. N., Rodd N. L., Slatyer T. R., 2016, Phys. Dark Univ., 12, 1
Di Mauro M., 2021, Phys. Rev. D, 103, 063029
Dinsmore J. T., Slatyer T. R., 2022, JCAP, 2022, 025
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Freudenreich H. T., 1998, ApJ, 492, 495 (F98)
Gautam A., Crocker R. M., Ferrario L., Ruiter A. J., Ploeg H., Gordon C., Macias O., 2022, Nat. Astron., 6, 703
Goodenough L., Hooper D., 2009, preprint (arXiv:0910.2998)
Gordon C., Macias O., 2013, Phys. Rev. D, 88, 083521
Grand R. J. J., White S. D. M., 2022, MNRAS, 511, L55
Haggard D., Heinke C., Linden T., Hooper D., Linden T., 2017, J. Cosmol. Astropart. Phys., 2017, 0
Hooper D., Goodenough L., 2011, Phys. Lett. B, 697, 412
Hooper D., Linden T., 2011, Phys. Rev. D, 84, 123005
Hooper D., Mohlabeng G., 2016, J. Cosmol. Astropart. Phys., 2016, 049
Hooper D., Slatyer T. R., 2013, Phys. Dark Univ., 2, 118
Launhardt R., Zylka R., Mezger P. G., 2002, A&A, 384, 112
Lazar A. et al., 2020, MNRAS, 497, 2393
Leane R. K., Slatyer T. R., 2019, Phys. Rev. Lett., 123, 241101
Leane R. K., Slatyer T. R., 2020a, Phys. Rev. D, 102, 063019
Leane R. K., Slatyer T. R., 2020b, Phys. Rev. Lett., 125, 121105
Lee S. K., Lisanti M., Safdi B. R., Slatyer T. R., Xue W., 2016, Phys. Rev. Lett., 116, 051103
Linden T., Rodd N. L., Safdi B. R., Slatyer T. R., 2016, Phys. Rev. D, 94, 103013
List F., Rodd N. L., Lewis G. F., Bhat I., 2020, Phys. Rev. Lett., 125, 241102
List F., Rodd N. L., Lewis G. F., 2021, Phys. Rev. D, 104, 123022
Macias O., Gordon C., Crocker R. M., Coleman B., Paterson D., Horiuchi S., Pohl M., 2018, Nat. Astron., 2, 387
Macias O., Horiuchi S., Kaplinghat M., Gordon C., Crocker R. M., Nataf D. M., 2019, J. Cosmol. Astropart. Phys., 2019, 042
Macias O., van Leijen H., Song D., Ando S., Horiuchi S., Crocker R. M., 2021, MNRAS, 506, 1741
Malyshev D. V., 2024, preprint (arXiv:2401.04565)
Manconi S., Calore F., Donato F., 2024, preprint (arXiv:2402.04733)
McCann A., 2015, ApJ, 804, 86
McDermott S. D., Zhong Y.-M., Cholis I., 2023, MNRAS, 522, L21(M2023)
Mishra-Sharma S., Cranmer K., 2022, Phys. Rev. D, 105, 063017
Nishiyama S. et al., 2013, ApJ, 769, L28
Ploeg H., Gordon C., Crocker R., Macias O., 2020, JCAP, 2020, 035
Pohl M., Macias O., Coleman P., Gordon C., 2022, ApJ, 929, 136
Porter T. A., Johannesson G., Moskalenko I. V., 2017, ApJ, 846, 67
Song D., Macias O., Horiuchi S., 2019, Phys. Rev. D, 99, 123020
Storm E., Weniger C., Calore F., 2017, J. Cosmol. Astropart. Phys., 2017, 022
Strong A. W., Moskalenko I. V., 1998, ApJ, 509, 212
Trotta R., 2008, Contemp. Phys., 49, 71
Vitale V., Morselli A., 2009, preprint (arXiv:0912.3828)
Weiland J. L., Arendt R. G., Berriman G. B., 1994, ApJ, 425, L81
Wolleben M., 2007, ApJ, 664, 349
Zhong Y.-M., Cholis I., 2024, preprint (arXiv:2401.02481)
Zhou B., Liang Y.-F., Huang X., Li X., Fan Y.-Z., Feng L., Chang J., 2015, Phys. Rev. D, 91, 123010

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