The Andromeda gamma-ray excess

background systematics of the millisecond pulsars and dark matter interpretations

Zimmer, F.; Maclas, O.; Ando, Shin'ichiro; Crocker, R.M.; Horiuchi, S.

DOI
10.1093/mnras/stac2464

Publication date
2022

Document Version
Final published version

Published in
Monthly Notices of the Royal Astronomical Society

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **516,** 4469–4483 (2022) https://doi.org/10.1093/mnras/stac2464
Advance Access publication 2022 September 8

The Andromeda gamma-ray excess: background systematics of the millisecond pulsars and dark matter interpretations

Fabian Zimmer [1]★ Oscar Macias,[1,2]★ Shin'ichiro Ando,[1,2,3]★ Roland M. Crocker [4] and Shunsaku Horiuchi [3,5]

[1]*GRAPPA – Gravitational and Astroparticle Physics Amsterdam, University of Amsterdam, Science Park 904, NL-1098 XH Amsterdam, the Netherlands*
[2]*Institute for Theoretical Physics Amsterdam and Delta Institute for Theoretical Physics, University of Amsterdam, Science Park 904, NL-1098 XH Amsterdam, the Netherlands*
[3]*Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*
[4]*Research School of Astronomy and Astrophysics, Australian National University, Canberra 2611, A.C.T., Australia*
[5]*Center for Neutrino Physics, Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA*

Accepted 2022 August 25. Received 2022 July 28; in original form 2022 April 18

ABSTRACT

Since the discovery of an excess in gamma rays in the direction of M31, its cause has been unclear. Published interpretations focus on dark matter or stellar related origins. Studies of a similar excess in the Milky Way centre motivate a correlation of the spatial morphology of the signal with the distribution of stellar mass in M31. However, a robust determination of the best theory for the observed excess emission is challenging due to uncertainties in the astrophysical gamma-ray foreground model. We perform a spectro-morphological analysis of the M31 gamma-ray excess using state-of-the-art templates for the distribution of stellar mass in M31 and novel astrophysical foreground models for its sky region. We construct maps for the old stellar populations of M31 based on data from the PANDAS survey and carefully remove the foreground stars. We also produce improved astrophysical foreground models via novel image inpainting techniques based on machine learning methods. Our stellar maps, mimicking the location of a population of millisecond pulsars in the bulge of M31, reach a 5.4σ significance, making them as strongly favoured as the simple phenomenological models usually considered in the literature, e.g. disc-like templates. This detection is robust to generous variations of the astrophysical foreground model. Once the stellar templates are included in the astrophysical model, we show that the dark matter annihilation interpretation of the signal is unwarranted. We demonstrate that about one million unresolved millisecond pulsars naturally explain the observed gamma-ray luminosity per stellar mass, energy spectrum, and stellar bulge-to-disc flux ratio.

Key words: pulsars: general – dark matter – gamma-rays: diffuse background.

1 INTRODUCTION

Due to its proximity and mass, the centre of the Milky Way (MW) is expected to be the brightest source of dark matter (DM) annihilation in the sky (e.g. Bertone, Hooper & Silk 2005; Charles et al. 2016). However, our view of the MW halo is obscured by large amounts of uncertain interstellar material. It is thus vital to carry out complementary searches for DM emission in regions with differing astrophysical uncertainties.

The Andromeda galaxy (M31), located at a somewhat low Galactic latitude (l, b) = (121.17°, −21.57°), suffers lower foreground extinction than the MW centre. Moreover, due to the large inclination angle (∼77.5°) of the plane of M31 with the line of sight (Tamm et al. 2012), we can observe the M31's halo almost completely unobstructed by its stellar and gaseous disc. A wealth of recent observations serve to elucidate key differences between the MW and M31, which are thought to be due to their different accretion histories (Kormendy 2013). M31 has a stellar mass that is similar or possibly even larger than that of the MW (e.g. Watkins, Evans & An 2010; Diaz et al. 2014). Furthermore, the M31's bulge is a factor of 5 to 6 times more massive than the compact bulge/bar system in the MW (Tamm et al. 2012; Licquia & Newman 2015), its supermassive black hole is about 50 times more massive than Sgr A⋆ (Bland-Hawthorn & Gerhard 2016), and the star formation rate (SFR) in M31 is a factor of about 10 times lower than in the MW (Ford et al. 2013). Altogether, M31 constitutes not only an excellent target for searches of DM emission (e.g. Lisanti et al. 2018), but also a unique stepping stone in our efforts to understand the high-energy astrophysics of spiral galaxies.

Interestingly, there has been a recent discovery of an excess in gamma-rays coming from the inner region of M31 (Ackermann et al. 2017a), which appears to have similar characteristics to those of the long-standing so-called Galactic Centre Excess (GCE) in the MW (e.g. Hooper & Goodenough 2011; Abazajian & Kaplinghat 2012; Gordon & Macias 2013; Macias & Gordon 2014; Calore et al. 2015; Ajello et al. 2016; Daylan et al. 2016; Ackermann et al. 2017b).

★ E-mail: f.zimmer@uva.nl (FZ); o.a.maciasramirez@uva.nl (OM); s.ando@uva.nl (SA)

In particular, using almost 7 yr of *Fermi*-LAT data, Ackermann et al. (2017a) detected extended diffuse gamma-ray emission from the inner ∼0.4° (or ∼5 kpc) of the M31's halo, at the 4σ statistical level. Its spatial morphology was found to be compatible with a uniform/Gaussian disc, and its spectrum was not well constrained – either a simple power-law or a bump-like spectrum could give an acceptable fit to the data. Importantly, they obtained that M31's emission is not correlated with the distribution of interstellar gas or regions of star formation activity, which are both mostly located in a large ring-like structure at ∼10 kpc from its centre. However, gamma-ray emission from the gaseous disc was not strongly ruled out and might even be present up to the 50 per cent level of the measured flux (Ackermann et al. 2017a). In addition, a later in-depth large field study (Karwin et al. 2019) of the outer halo of M31 found evidence for emission within the ∼120–200 kpc (or ∼8°–12°) of its centre.

Recent theoretical studies have proposed different alternative explanations for the inner galaxy M31 excess: cosmic ray (CR) models (e.g. McDaniel, Jeltema & Profumo 2019; Do et al. 2021), unresolved millisecond pulsars (MSPs) (e.g. Eckner et al. 2018; Fragione, Antonini & Gnedin 2019), and DM emission models (e.g. Burns et al. 2021; Chan et al. 2021; Karwin et al. 2021). As for the excess in the outer halo of M31, Recchia et al. (2021) posited some well-motivated CR models.

Since the M31 signal is only an 'excess' with respect to our current understanding of the diffuse gamma-ray background, it is important to investigate the impact that diffuse mismodelling has on the characteristics of the signal. Very recent studies (e.g. Di Mauro et al. 2019; Armand & Calore 2021) have used either `SkyFACT` (Storm, Weniger & Calore 2017), or alternative diffuse emission models built with GALPROP (Strong & Moskalenko 1998) for the evaluation of these uncertainties. In this study, we perform a reanalysis of these data with a novel approach; we construct tailor-made foreground gas map models for the M31 region which are expected to be less affected by biases. Karwin et al. (2021) highlighted that the foreground interstellar gas maps used in the identification of the M31 excess might contain a fraction of the interstellar gas that should belong to the M31 galaxy. Here, we excise M31 from the foreground interstellar gas models, followed by using various inpainting techniques (e.g. the methods 'Nearest-Neighbour' and '`SMILE`') for reconstruction or processing all the information of the image by a neural network ('Deep-Prior' method) to inpaint the excised region. In this way, we restore the actual spatial distribution of the interstellar gas in the foreground as best as possible. The resulting alternative maps are then employed in our reassessment of the characteristics of the M31 excess.

Studies of the GCE (Bartels et al. 2018; Macias et al. 2018, 2019; Abazajian et al. 2020; Coleman et al. 2020; Pohl et al. 2022) have demonstrated that the two most compelling explanations (astrophysical or DM emission) can be distinguished based on their spatial morphologies. Indeed, such studies have consistently found that the morphology of the GCE traces the stars of the inner Galaxy better than it traces the distribution expected for DM (however, see Di Mauro 2021) for opposite conclusions). Motivated by the very promising morphological results obtained in the GCE, here we attempt to separate the stellar from the DM hypotheses using a traditional fitting regression technique. For this, we construct empirical models of the stellar distribution in M31 using data from the PANDAS survey, which has been mapping out the environment of M31 for more than a decade. We construct stellar maps for the disc based on these data and using the stellar contamination model of Martin et al. (2013), which effectively removes the stars residing in the MW. Since, unfortunately, these data are unreliable for the innermost regions of M31, we use instead data from WISE (Wright et al. 2010) to construct the bulge component of our stellar templates.

The paper is outlined as follows. In Section 2, we describe the data used in this work. In Section 3, we present the methods used to construct the diffuse emission components and the various models used in our analysis. In Section 4, we present the results of the gamma-ray analyses and discuss their implications in Section 5.

2 GAMMA-RAY OBSERVATIONS

In this work, we use data from the *Fermi* Large Area Telescope (*Fermi*-LAT), covering 10 yr of observations (see Table 1). We use exactly the same data selection cuts assumed in the construction of the 4FGL-DR2 catalogue (Abdollahi et al. 2020; Ballet et al. 2020) to avoid potential biases induced by new point sources (not present in the 4FGL-DR2 catalogue) which would likely appear in our region of interest (ROI) if we were using a longer observation time. Additionally, we have followed the recommendations of the *Fermi* collaboration[1] by choosing to set the maximum zenith angle at $z_{\rm max} = 90°$, appropriate for our type of analysis and data.

We restrict ourselves to data from a 14 × 14 deg region in the sky, centred on the SIMBAD coordinates of M31, i.e. $(l, b) = (121.17°, -21.57°)$, and spanning an energy range of 500 MeV to 100 GeV. Although previous studies have gone lower to 300 MeV (Armand & Calore 2021) or even as low as 100 MeV (Di Mauro et al. 2019), we aim to preserve as much angular information as possible, without too much of a loss of photon statistics. We are most interested in the spatial morphology of the signal and we want to determine in which part of the M31 galaxy the signal is strongest. Since the containment angle of the *Fermi*-LAT quickly rises from ∼1° at 500 MeV to ∼5° at 100 MeV, greatly hindering the realization of our main goal, we have opted for the lower limit of said 500 MeV.

We use the *Fermi* `SCIENCETOOLS` version 2.0.0 software package to further reduce the raw data. Since the Pass 8 data release, it is possible to select data based on the reliability of the reconstructed direction based on the instrument's point spread function (PSF). We want to use photons from the PSF3 quartile only (the most reliable in reconstructed direction), which would benefit our morphological analysis the most. However, due to our bin-by-bin approach, higher energy bins lack statistical power if we do not include other PSF quartiles causing reconstructed spectra to fluctuate. Therefore, we restricted our analysis to the PSF3 quartile only for the first three energy bins, where the counts were high enough, and we used all PSF quartiles combined for the rest of the bins.

We restrict ourselves to the *CLEAN* event class data. There are other event classes, which are also appropriate for our kind of analysis. However, the majority of photon counts is contained in our first few energy bins and other possible choices, such as e.g. *SOURCEVETO*, have an identical background rate in this energy range and make for equally good choices. Ultimately, the choice of the *CLEAN* event class data is based on the fact that it constitutes a good compromise between photon statistics and the reliability of their astrophysical origin.

All the technical details of the selection filters applied to the data by the Fermitools are summarized in Table 1. For more information regarding the criteria, we refer the reader to the official *Fermi* website.[2]

[1] Zenith angle cut recommendations.
[2] *Fermi*-LAT Cicerone

Table 1. *Fermi*-LAT data selection criteria used for our analysis (applied to the data with the *Fermi* `ScienceTools`).

Category	Selection criteria
Observation period	2008 Aug 4 to 2018 Aug 2
Mission elapsed time (s)	239 557 417 to 554 929 985
Energy range (GeV)	0.5 to 100
Energy binning	10 bins (log spaced)
Event class	Pass 8 CLEAN
Event type	PSF3 for [0.5,2.5] GeV, all PSF for [2.5,100] GeV
IRFS	P8R3 CLEAN V3
$z_{\rm max}$	90°
Filters	(DATA_QUAL>0) & (LAT_CONFIG=1)
Region of interest	$14^\circ \times 14^\circ$ at $(l, b) = (121.17^\circ, -21.57^\circ)$
Pixel resolution	0.1°
Pixel binning	140×140 pixels
Sky projection	Cartesian 'CAR'

3 METHODS

In this section, we describe the construction of tailor-made Galactic diffuse emission models for our ROI. One of the biggest challenges in creating an appropriate foreground model for the M31 region is that it is very difficult to disentangle the foreground hydrogen gas that belongs to the MW from the actual gas in M31. The standard approach followed by the *Fermi* team consists of removing M31 from the hydrogen gas maps, based on cuts in the (l, b, v_{LSR}) data space with v_{LSR} being the local standard of rest velocity. After it has been excised, they reconstruct the foreground hydrogen maps based on the techniques described in Elad et al. (2005) and Acero et al. (2016a). Although this is a reasonable analysis choice, it is expected to introduce a bias in the gamma-ray properties of M31. In particular, it has been pointed out (Karwin et al. 2019) that the hydrogen maps used in the official foreground *Fermi* Galactic diffuse emission map might be holding some hydrogen gas which belongs to M31. In this work, we aim to construct more suitable Galactic foreground maps by using novel inpainting algorithms based on neural networks and deep learning to construct M31-free Galactic foreground templates. We also evaluate the uncertainties associated with this method by rerunning our gamma-ray pipeline with all the resulting alternative foreground maps.

3.1 Standard *Fermi* Galactic diffuse emission model

One of the models we considered in this work is the so-called `gll_iem_v07` model. This is the standard Galactic diffuse emission model recommended for analyses of point-like and small-sized extended sources, constructed by the *Fermi* collaboration – henceforth simply called the *Fermi* background.

The *Fermi* background is made up of a linear combination of templates, each one responsible for capturing the contribution of a physical mechanism to the total gamma-ray emission. Some of the largest contributions come from the decay of energetic neutral pions, and electron bremsstrahlung radiation. Since these two components are spatially correlated with gas, they were phenomenologically modelled using atomic and molecular hydrogen maps constructed from radio observations (Acero et al. 2016a).

Another big contribution to the *Fermi* background comes from the inverse-Compton scattering process. This component is very difficult to model correctly as it depends on many factors, which include solving the cosmic ray transport equation. The *Fermi* team included inverse-Compton emission templates generated with the numerical code `GALPROP` (Strong & Moskalenko 1998).

The procedure followed by the *Fermi* team was then to fit these components – in addition to other empirical templates accounting for positive and negative residuals in the data – to the observed all-sky gamma-ray data using a template fitting approach. Once the best-fitting fluxes for each template were determined, they combined all these in one compactified energy-dependent spatial template. This constitutes the `gll_iem_v07` model.

3.2 Alternative foreground/background diffuse emission models

3.2.1 *Gas-correlated gamma-ray component*

As in the *Fermi* background, the gas-correlated emission is modelled by assuming that the gamma-ray intensity is proportional to the interstellar gas column density – mainly atomic (HI) and molecular (H2) hydrogen. We use the high-resolution HI4PI survey (HI4PI Collaboration 2016) as in the *Fermi* background as our first option. As a second option we use hydrodynamical gas and dust maps, i.e. HI and H2 gas column density templates developed in Macias et al. (2018), which were already used to show that the GCE can be explained by an unresolved MSP population rather than annihilating DM (Macias et al. 2019; Abazajian et al. 2020). These templates were originally designed for analyses related to the MW Galactic centre and for this purpose split into four Galactocentric rings. Only the rings 3 and 4 of the HI component contribute to the M31 region, whereas none do for the H2 component.

Similarly to the *Fermi* collaboration, we also excise M31 from the gas maps where present, but use more sophisticated inpainting algorithms to fill in the regions of missing data. Two methods for inpainting are employed. The first one uses deep convolutional neural networks to restore the image. It was developed in Puglisi & Bai (2020), where they successfully recovered the necessary statistical properties the image had before the reconstruction process. The authors tested it in the context of synchrotron and dust polarized emission, which represent the Galactic contamination in cosmic microwave background measurements. Their code was made publicly available as the **P**ython **I**npainter for **C**osmological and **AS**trophysical **SO**urces (`PICASSO`) package. We successfully employed two methods of their code for our purposes: The 'Nearest-Neighbour' algorithm based on Bucher, Racine & van Tent (2016) and the 'Deep-Prior' method based on Ulyanov, Vedaldi & Lempitsky (2020).

The other inpainting algorithm we use is from Li (2014), which we call the `SMILE` inpainting algorithm. It is based on solving the Laplace equation for each missing pixel by using the values of the surrounding pixels.[3] Such an approach has been successfully used by Macias et al. (2019) to inpaint masked point-like sources near the centre of the MW.

Visual representations of the inpainting procedures when using either `SMILE` or `PICASSO` are shown in Figs 1 and 2, respectively. The first column of panels in each figure show the original map, where M31 is present. The second column depicts the extension of the masked region, which we based on cuts described in Ackermann et al. (2017a), which were shown to effectively remove the majority

[3]For an example of a similar method implemented in Mathematica see https://community.wolfram.com/groups/-/m/t/873396

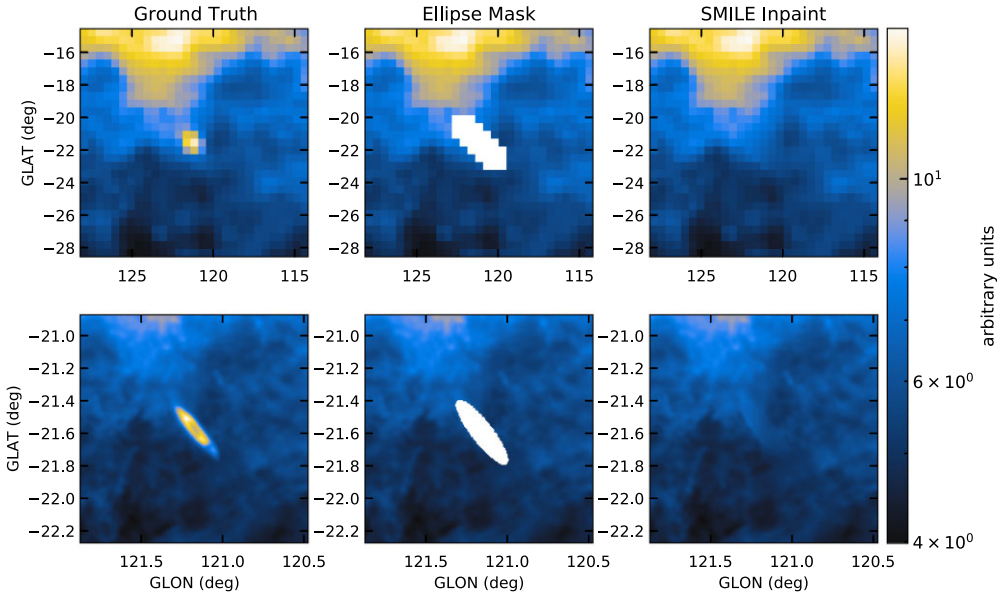

Figure 1. Visualization of the inpainting process when using the SMILE algorithm. From left to right, the three columns show the original hydrogen template, the masked region, and the inpainted final version. The upper and lower rows show the HI component of Macias et al. (2018) (as described in the text) and the HI4PI map, respectively.

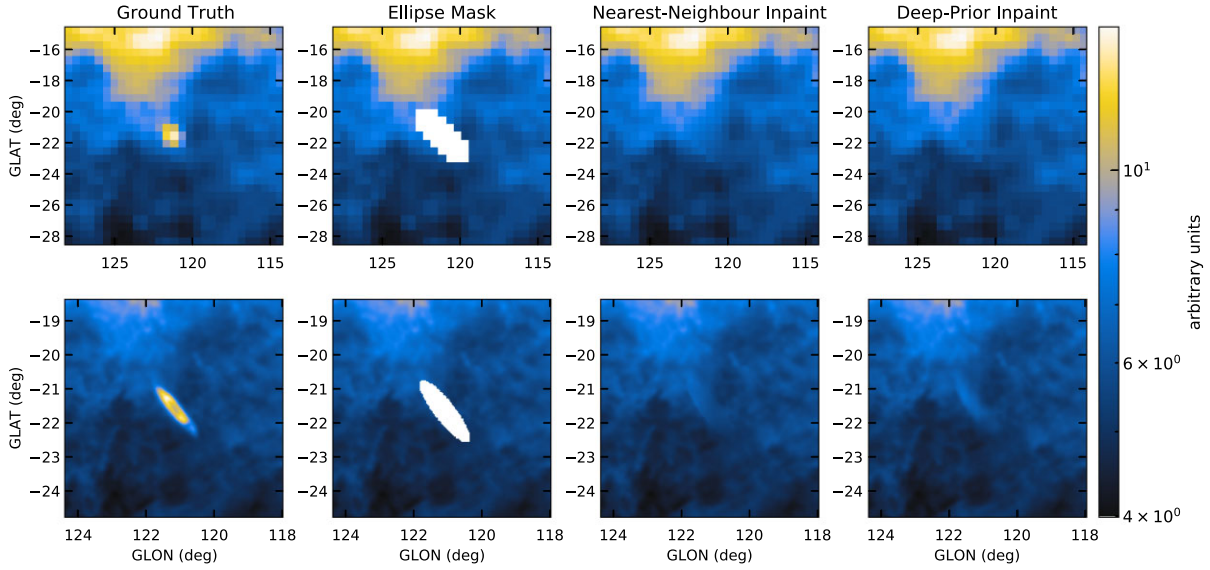

Figure 2. Visualization of the inpainting process when using the PICASSO algorithms. From left to right, the four columns show the original hydrogen template, the masked region, and the final inpainted versions when using the Nearest-Neighbour or the Deep-Prior algorithm. The upper and lower rows show the HI component of Macias et al. (2018) (as described in the text) and the HI4PI map, respectively.

of the disc of M31 from the data. The region was then inpainted with the aforementioned tools and we chose to use these algorithms over the one previously used by the *Fermi* collaboration due to their simpler implementation and to assess the impact on the properties of the M31 excess due to systematic uncertainties in the gamma-ray background model. These inpainted, M31-free versions can be seen in the last columns. In both figures, the lower row corresponds to the HI4PI map and the upper row to the atomic hydrogen column density map used in Macias et al. (2018). Note that the latter was derived from the 21 cm LAB survey (Kalberla et al. 2005), which has much lower angular resolution than the newer HI4PI survey.

3.2.2 *Inverse-compton component*

The second largest contribution to the diffuse gamma-ray sky comes from the inverse-Compton radiation produced by the photon upscattering by energetic electrons. All of the models for this inverse-Compton component we employ in this work are generated with the Galactic cosmic ray propagation code GALPROP (Strong & Moskalenko 1998).

We used the three models from Ackermann et al. (2015) labelled Foreground Model A, B, and C. This study was interested in investigating how their fit results were affected by the specific type

of foreground model chosen. These three models are thought to encompass a very wide range of systematic uncertainties associated with the inverse-Compton component. Model A uses a distribution of cosmic ray sources based on Lorimer et al. (2006) and standard choices for the propagation parameter set-up. For Model B they include an additional population of sources for electrons near the Galactic centre, which made this model a better fit according to their analysis. Also, Model B predicts an enhanced gamma-ray emission at high latitudes. This makes it highly relevant for our study as our ROI is relatively far from the plane. The diffusion and reacceleration parameters of Model C vary with Galactocentric radius throughout the Galaxy, as opposed to the constant diffusion coefficient assumed in the two other models.

Additionally, we use the inverse-Compton model of Abazajian et al. (2020), which is split into six Galactocentric rings. Only their local and outermost rings contribute to the M31 sky region and are used in our analyses.

By combining templates for these two major components and different inpainting algorithms we arrive at a total of 2 (H1 components) × 3 (Inpainting Methods) × 4 (Inverse-Compton Models) = 24 combinations. By testing models over this kind of variety of background models, we can classify the obtained significances of detection against the systematic uncertainties that come with the modelling of the Galactic diffuse gamma-ray emission. In addition to the gas-correlated and inverse-Compton templates, we use the same components as in the *Fermi* background for the isotropic gamma-ray background, emission from the sun and moon, and the same 4FGL-DR2 source catalogue.

3.3 Phenomenological M31 models

All the phenomenological models we used are integrated in `Fermitools`. We tested whether the signal was point-like or spatially extended. Specifically, we used discs of different radii and uniform brightness or a Gaussian disc of different variance. In most of the studies related to this phenomenon, investigators have included similar templates in their analyses, from the discovery of the excess (Ackermann et al. 2017a) to both the previously mentioned studies closely related to our work (Di Mauro et al. 2019; Armand & Calore 2021). The purpose of these simple constructions is to test whether the data prefer an extension beyond a point-like emission, but these extended templates are likely unphysical in the sense that real emission will fail to be as symmetric and clean as they are. Arguably, better suited models for this type of analysis are based on multiwavelength observations from space telescopes and ground detectors.

3.4 Interstellar gas and dust models of M31

One of the difficulties in astronomical studies like ours, where we want to isolate specific astrophysical gamma-ray sources, is the interference of emission from gas and dust clouds. They can either be part of the Andromeda system or reside inside the MW halo, extending across the line of sight between us and M31. There has even been evidence for hydrogen structures extending between the two galaxies in interstellar space (Lockman, Free & Shields 2012).

There have been studies dedicated to mapping out the gas and dust contents of the Andromeda galaxy. A deep wide-field HI imaging survey has been done in Braun et al. (2009) with the *Westerbork Synthesis Radio Telescope*. We also include two different versions of this hydrogen map labelled by BraunV2 and BraunV3, which include corrections for opacity effects. An intensive study about Andromeda's dust was done in Draine et al. (2014) using observations from the *Spitzer Space Telescope* and *Herschel Space Observatory*. We use these three maps[4] to account for unmodelled residual gas and dust emission belonging to M31.

3.5 Construction of the stellar density templates

One of the most important undertakings of this work is to test whether the data prefer an explanation involving MSPs or DM. To that end, we constructed stellar density maps consisting of old red giants (from now on simply called stellar maps), which we use as proxy for the spatial distribution of MSPs given that they are both old stellar populations.

We will begin by describing the construction of our stellar maps, for which we adapted the procedure of Martin et al. (2013). The data used to construct our stellar maps stem from the Pan-Andromeda Archaeological Survey (PAndAS) (Ibata et al. 2014; McConnachie et al. 2018). The data are publicly available at the PAndAS archive.[5]

First, we selected all stars within a 150 kpc radius from the PAndAS survey, centred on the SIMBAD coordinates of the centre of M31. These stars are then de-reddened to correct for extinction by using the $E(B-V)$ map from Schlegel, Finkbeiner & Davis (1998) and the corrected colour magnitudes are then obtained with

$$\begin{aligned} g_0 &= g - 3.793E(B-V) \\ i_0 &= i - 2.086E(B-V). \end{aligned} \tag{1}$$

The observed magnitudes as stated in the PAndAS survey are g and i and their de-reddened equivalent have the $_0$ subscript, which we omit from now on. In the colour–magnitude space of all these stars, we selected only a subsample corresponding to stars which have colour and magnitude as expected from old, low-metallicity stars. This 'selection box', as depicted in Fig. 3, is based on isochrones of a stellar population with a certain age and metallicity. For our purpose, we generated[6] isochrones for six different age populations (3, 5, 7, 9, 11, and 13 Gyr), each having five different metallicities, i.e. −[Fe/H] ratios (0.5, 1.0, 1.5, 2.0, and 2.5). They were adjusted for the M31 distance modulus based on Conn et al. (2011, 2012).

The data of the PAndAS survey have holes in some places due to gaps in the coverage, saturated bright stars, or instrumental failures. We reconstructed these missing data by using our `SMILE` inpainting algorithm.

Since we are ultimately interested in the old stars, which are part of the Andromeda system, the stars residing in the halo of the MW are therefore contaminants and have to be removed. The density of these contamination stars, at a specific colour and magnitude $(g-i, i)$, is modelled as an exponential increase towards the Galactic plane and centre of the MW as

$$\Sigma_{(g-i,i)}(X,Y) = exp(\alpha_{(g-i,i)}X + \beta_{(g-i,i)}Y + \gamma_{(g-i,i)}), \tag{2}$$

where the coordinates (X, Y) are the equatorial coordinates (α, δ), projected on a plane tangential to the M31 centred celestial sphere. The functions α, β, and γ[7] are constructed empirically from regions near M31, where contamination of the MW foreground stars is low. For more details regarding these functions, see Martin et al. (2013).

[4]All data products related to these studies can be downloaded from https://www.astro.princeton.edu/~draine/m31dust/m31dust.html.
[5]Queries can be made here.
[6]We used isochrones from Dotter et al. (2008), which can be generated at the Dartmouth Stellar Evolution Database.
[7]These functions were kindly provided to us by the authors of the paper.

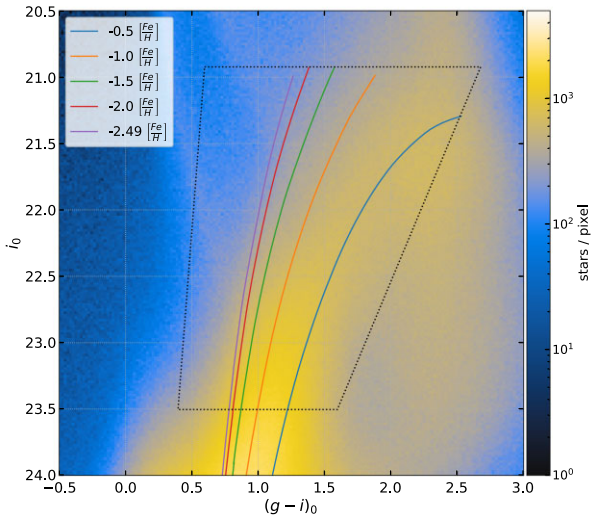

Figure 3. Colour–magnitude diagram of all stars in a 150 kpc radius of M31 from the PAndAS survey with a 0.02 × 0.02 magnitude pixel size. The selection box boundaries are the same as in Martin et al. (2013), where the upper and right box limit is determined by the tip of the red giant branch (the highest reaching isochrone) and the rightmost reaching isochrone (the reddest star), respectively.

A visualization of the individual steps of the construction for these maps can be seen in Fig. 4.

This procedure gives us a stellar map for each of our six selected age populations ranging from 3 to 13 Gyr. These are now subtracted from each other in such a way that we get five stellar maps containing 3–5, 5–7, 7–11, and 11–13 Gyr population groups.

Unfortunately, the data of this survey are not reliable in the bulge area of M31. This is mostly due to the telescope not being able to resolve individual stars in this bright region of the sky. We therefore follow the common procedure (see e.g. Martin et al. 2014) of masking the bulge area in these maps and instead use data from the Wide-field Infrared Survey Explorer (WISE) (Wright et al. 2010) at wavelengths of 3.4 (W1) and 4.6 (W2) microns,[8] which was shown to be best suited for tracing stellar light rather than dust (e.g. Ness & Lang 2016).

The dimension of this mask is based on a study of the structural parameters of M31 (Courteau et al. 2011), which contains estimates of the extension of the bulge based on its luminosity profile (see their fig. 9). We show a close-up of this masked region together with the stellar density profile along the major and minor axis of M31 in Fig. 5. Note that using the raw W1 and W2 templates away from the bulge is not well justified given that these are expected to be heavily contaminated by dust/gas emission from the disc. In contrast to the MW centre, we observe the bulge of M31 almost completely unobstructed by its own gaseous disc, due to the large inclination angle of M31. This further supports the assumption that the infrared W1/W2 data in the bulge are mostly produced by the stars.

One of the key differences in this work compared to previous works is the use of empirical data over smooth analytical functions for our stellar templates. This guarantees a more sensitive comparison of the MSP to the DM hypotheses in our later analysis. A different approach based on using Einasto density profiles to model the bulge and disc of M31 has been used in e.g. Armand & Calore (2021). We compare their bulge Einasto profile to the radial density profile of our stellar map for W1 and W2 in the top and bottom row of Fig. 6, respectively. The middle and right-hand panel show the normalized density along the major and minor axis, seen as black dotted lines in the left-hand panel, respectively. Additionally, we have added a squared NFW density profile, representing our DM template, demonstrating that it is more concentrated to the innermost regions of M31 compared to the other two profiles, which impacts the results of our morphological analysis.

The combination of these masked stellar maps (each containing a different age group) with either a W1 or W2 bulge acts as our final model for the MSPs hypothesis.

3.6 DM model

Our DM template has to account for both the M31 and MW DM haloes, since our line of sight extends through both of them. We model the radial density profiles of both haloes with the symmetric NFW profile (Navarro, Frenk & White 1996) given by

$$\rho(r) = \frac{\rho_0}{\frac{r}{r_s}(1+\frac{r}{r_s})^2}. \tag{3}$$

For the density ρ_0 and the scale radius r_s for M31 and the MW we use the parameters from Karwin et al. (2019). We begin by expressing the radius from the centre of either Andromeda or our Galaxy in terms of the line-of-sight variable s and the Galactic longitude and latitude coordinates l and b (centred on M31's SIMBAD coordinates in case of the M31 DM halo) with the law of cosines as

$$r = \sqrt{d^2 + s^2 - 2ds\cos(l)\cos(b)}. \tag{4}$$

To obtain the value for one pixel of our template, the DM density is squared[9] and integrated along the line of sight as

$$J_p(l,b) = \int_s \rho(r[s,l,b])^2 \mathrm{d}s. \tag{5}$$

The total value of each pixel J_p, also referred to as the J-factor, is then the sum of the contributions of both DM haloes. This is repeated for each pixel to obtain the final template. Fig. 7 shows the J-factor profile as a function of distance from the M31's centre. It is clear that the J-factor contribution from the MW halo becomes dominant for angular distances $\gtrsim 1^\circ$ away from the centroid of M31 (Karwin et al. 2021). We also display (vertical lines) the size of our ROI in comparison with the extent of the observed gamma-ray excess, which demonstrates the adequacy of our analysis region. Furthermore, we compare our resulting radial J-factor profile with the one obtained by Karwin et al. (2021) (c.f. their fig. 6 and our Fig. 7) and find our values to be consistent with theirs.

3.7 Fitting procedure

Since the flux ratios of the components included in the *Fermi* background model were obtained from an all-sky fit, these do not necessarily represent well the flux ratios in a small patch of the sky, such as the 14 × 14 deg region around M31 we have in this work. Another important issue is that this compactified map does not permit a rigorous evaluation of the systematic uncertainties associated with the Galactic diffuse emission model.

In this work, we investigate the impact of each component making up the Galactic diffuse emission. For this, we fit all components of

[8]We used data available at NASA's SkyView.

[9]The flux from dark matter annihilation is proportional to the density profile squared, as the interaction requires two particles.

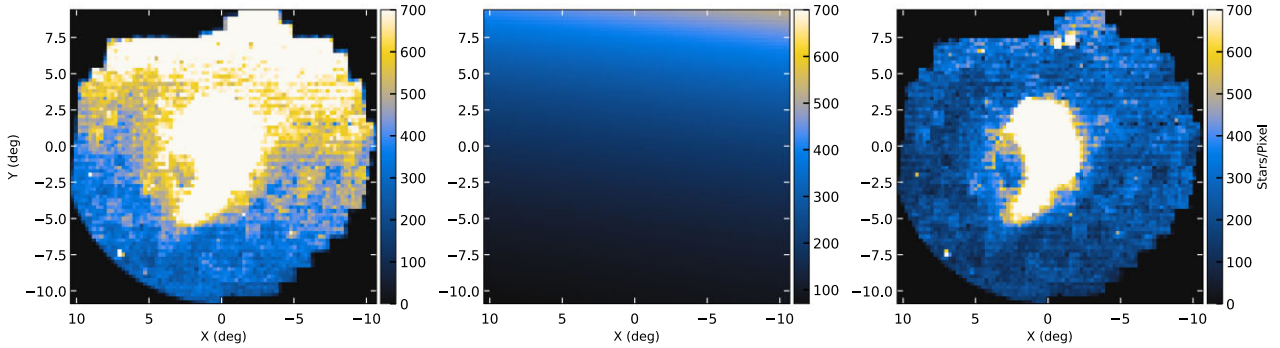

Figure 4. Visualization of the three stages of the stellar map construction. The coordinates (X, Y) are the equatorial coordinates (α, δ), projected on a plane tangential to the M31 centred celestial sphere. Left-hand panel: All stars in our line of sight, where some are part of the MW system. Middle panel: The model for the stars belonging to the MW system, which are contaminating the sample we are interested in, i.e. all stars of the M31 system. Right-hand panel: Residual stars after subtracting the contaminant stars, resulting in stars mostly belonging to Andromeda.

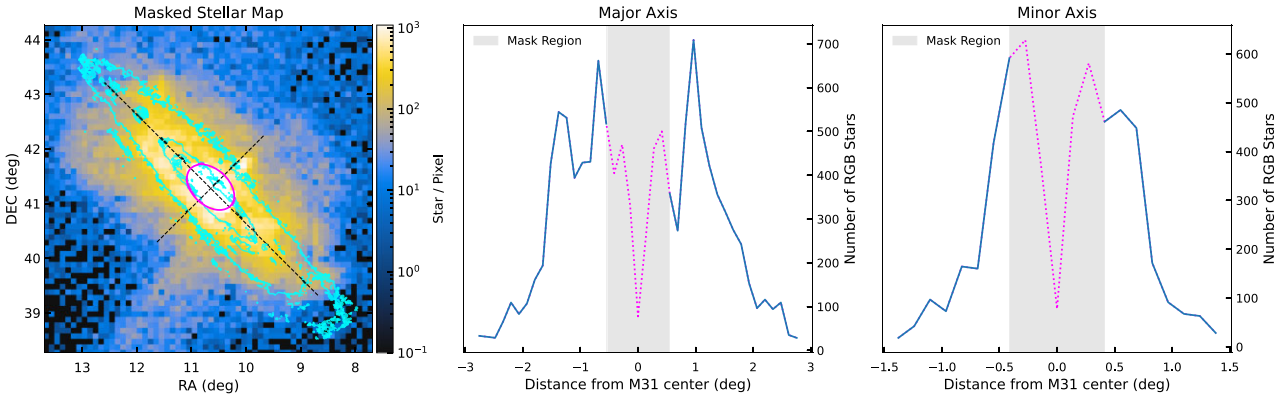

Figure 5. The left-hand panel shows the extent of the mask covering the centralized region, i.e. the bulge of our constructed stellar maps (for the 3–5 Gyr age map in this case). The cyan curves are gas contours to show the extent of the disc compared to the bulge of M31. The density of stars along the major and minor axis (the dotted black lines in the left-hand panel) of the elliptical shaped disc of Andromeda is shown in the middle and right-hand panel, respectively. The masked region is indicated by the grey shaded area, where we can see the decrease in the number of stars by the dotted magenta lines.

our alternative foreground emission model (FEM), as summarized in Table 2, individually using a bin-by-bin analysis procedure. We run a separate maximum-likelihood fit in each of these bins, where we use a simple power law ($\mathrm{d}N/\mathrm{d}E = N_0 E^{-\alpha}$) to model the spectra of every source in our ROI. We allowed the normalization N_0 of each source to vary, such that they have freedom to absorb any potential excess that is due to Galactic foreground emission. However, the spectral slope α was set to 2 to stay agnostic to the overall unknown spectral shape. The fits were performed with the `PyLikelihood` tool, where a total of 79 point sources were present in our ROI. The obtained likelihoods were used to compute the statistical significance of the various hypotheses in σ units through equations (A1) and (A4).

4 RESULTS

The contents of the previous sections can be summarized as follows. There are certain challenges we face when dealing with the complex nature of gamma-ray emission from M31. The main challenge is how to disentangle the foreground emission due to the MW from the emission in M31. We do so by constructing tailor-made emission templates for our ROI with the use of state-of-the-art image restoration techniques. The alternative M31-free foreground templates are used in this section to estimate the effects that systematic uncertainties on this component have on the properties of the M31 excess.

In summary, the foreground components comprising our baseline (also referred to as standard) model for the fits consist of the *Fermi* diffuse emission model described in Section 3.1, the extragalactic diffuse emission model `iso_P8R3_CLEAN_V3_PSF3_v1` for the first three energy bins, and `iso_P8R3_CLEAN_V3_v1` for the rest, together with the 79 point sources in our ROI from the 4FGL-DR2 source catalogue. Our alternative foreground model consists of either the HI4PI survey or the Galactocentric HI and H2 ring templates for the gas-correlated emission, and the models labelled A, B, or C for the inverse-Compton emission. The components modelling the extragalactic emission and γ-ray point sources are the same as in the baseline model. A list of all the templates considered in this work is shown in Table 2.

4.1 Spatial morphology of the M31 excess

4.1.1 Phenomenological models

We started our analysis runs with the phenomenological models (see Section 3.3). We are interested in comparing our findings to previous work on this matter, since we are using the CLEAN class data compared to the more commonly used SOURCE class data, which might influence the results of a morphological analysis significantly.

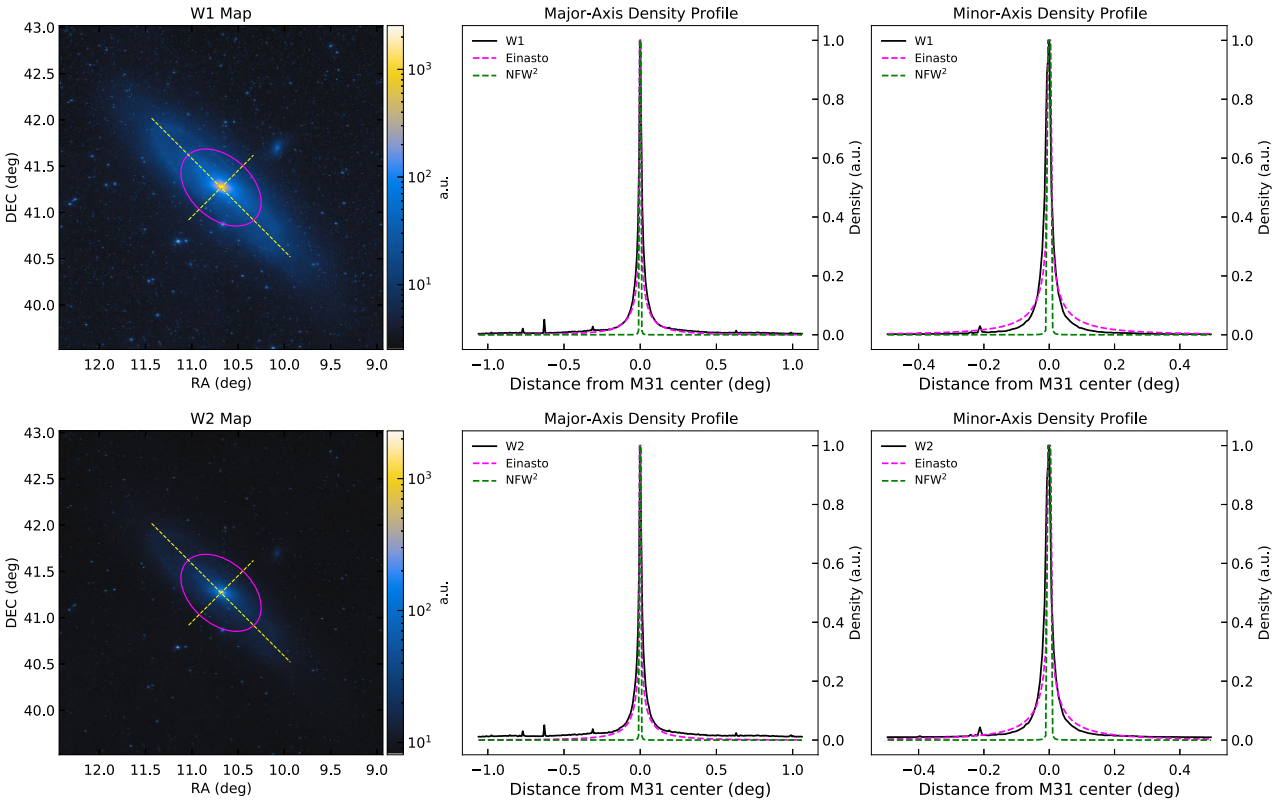

Figure 6. Radial density profile comparison of our W1 and W2 bulge templates along the major and minor axis (the dotted black lines in the left-hand panel) with the bulge Einasto profile used in Armand & Calore (2021) and the squared NFW profile. All the panels are shown in arbitrary units (a.u).

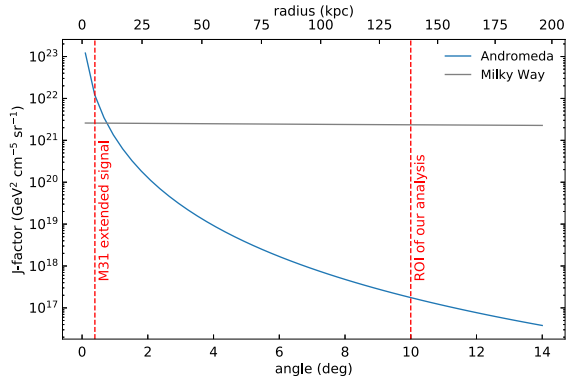

Figure 7. Radial profiles of the J-factor of our DM template for both the M31 and MW halo component, consistent with those of Karwin et al. (2021).

We tested two cases where the centre of the templates is either fixed or free. For the latter we moved the templates through a grid of locations. First, we moved it through a low-resolution grid with a stepsize of 0.1 deg in (RA, Dec.) space. We then followed up with a high-resolution grid with a stepsize of 0.01 deg in (RA, Dec.) space starting from the best-fitting spatial position of the low-resolution analysis. This procedure did not improve the fit much and yielded an almost identical value to the templates with a fixed centre, all lying in the range of 5.42σ to 5.46σ. All the results are summarized in Table 3.

To visualize the extension of these freed models, we superimpose them on a significance map of the M31 signal in Fig. 8(a). This significance map was obtained by moving a putative point-like source through a spatial coordinate grid and calculating the TS value at that point. The model used to test for the likelihood did not include a model for M31. The resulting map traces the morphology of the gamma-ray emission and is solely meant for visualizing the extent of the emission given our data.

A quick but important check is to see if there is anything we missed to model in our ROI, such as a point-like source we neglected to include in our background model. To see this, we again generate a significance map but this time we include M31 as modelled in the 4FGL-DR2 catalogue. We do not find any significant unmodelled excess in our ROI as can be seen in Fig. 8(b).

4.1.2 *Physically motivated models*

For our more physically motivated models (such as e.g. our stellar templates based on empirical observations), we first establish a significance ranking, for which we test each of our templates individually. We present the results of these tests in the upper section of Table 4. We found that our gas and dust templates all yielded similar results in the range from about 2.7σ to 3.0σ, showing that the M31 excess is not obviously correlated with emission from gas or dust. We also fitted a 150 MHz radio template, motivated by a study investigating the contribution of MSPs to this radio frequency (Sudoh, Linden & Beacom 2021), but could not obtain a significantly enough detection for that template with only $\sim 1.7\sigma$. Also, the signal does not seem to come from the disc alone, as our stellar disc templates all have rather low significances between $\sim 2.2\sigma$ and $\sim 2.4\sigma$. Most importantly, our results support previous works, finding the signal to be spatially correlated with the bulge of

Table 2. Summary of individual components comprising our baseline (also referred to as standard) and alternative foreground emission model used in our analysis. Each model component is described in Section 3.

Baseline foreground model components	**What γ-rays does it model?**
Fermi diffuse emission model	total Galactic diffuse emission
`iso_P8R3_CLEAN_V3_PSF3_v1`	extragalactic emission (first three energy bins)
`iso_P8R3_CLEAN_V3_v1`	extragalactic emission (all other energy bins)
4FGL-DR2	point sources
Alternative foreground model components	**What γ-rays does it model?**
HI4PI survey	gas-correlated emission
Galactocentric H*I* & H2 ring templates	gas-correlated emission
IC model A,B,C	Galactic inverse-Compton emission
`iso_P8R3_CLEAN_V3_PSF3_v1`	extragalactic emission (first three energy bins)
`iso_P8R3_CLEAN_V3_v1`	extragalactic emission (all other energy bins)
4FGL-DR2	point sources

Table 3. Significances for our phenomenological models for M31. The number of degrees of freedom of each template, i.e. the number of parameters left free to vary in the fit ($N_{d.o.f.}$), corresponds to 9.

M31 template	TS	σ	RA (deg)	Dec. (deg)	Radius (deg)
Point source (fixed)	41.83	4.92	10.6847	41.2687	–
Point source (free)	42.02	4.94	10.7047	41.2787	–
Uniform disc (fixed centre)	43.53	5.42	10.6847	41.2687	0.31
Uniform disc (free centre)	44.04	5.46	10.7347	41.2487	0.31
Gaussian disc (fixed centre)	43.46	5.42	10.6847	41.2687	0.16
Gaussian disc (free centre)	43.80	5.44	10.7347	41.2587	0.16

M31. In this ranking, our bulge templates even slightly outrank the DM template, with $\sim 5.4\sigma$ and $\sim 5\sigma$, respectively.

Since our stellar map is a two-component model comprised of a stellar disc component incorporating a certain red giant age population together with a component for the bulge based on WISE data, we show the significances for all 10 possible combinations of them. The results are displayed in the lower section of Table 4. We find that all combinations have a very similar significance of around 4.6σ. Naturally, we expect the MSPs to be present both in the disc and the bulge of M31. However, due to the weak nature of the excess and the lack of statistics, there is no strong support for both the stellar disc and bulge template together over the bulge template alone. However, we expect this to change with more future data and future γ-ray missions. We analyse the spectrum of the excess using this combination of templates rather than the bulge alone, since they differ by less than 1σ and the lack of statistics in the higher energy bins led to fluctuating results, when using the bulge template alone. Using the best-fitting 'disc+bulge' model, we display the residual photon counts map, i.e. the difference of the data to the model, in Fig. 9. In the left-hand panel, we only fit the *Fermi* foreground model to the data, whereas in the right-hand panel we added the disc+bulge model. The inclusion of this model has the effect of smoothing out the residuals, especially in the centre of the ROI, where the M31 excess is located.

4.2 Spectrum of the stellar disc plus bulge model

We also investigate the spectrum of our best-fitting stellar disc + bulge model, i.e. the model for the MSP population in the disc and bulge of M31. The spectrum together with previous results of the *Fermi* collaboration is shown in Fig. 10. The flux values are comparable to the study first documenting the Andromeda excess (Ackermann et al. 2017a) and to one of the most recent studies (Armand & Calore 2021) ensuring the validity of our results.

To accomplish our final goal, which is obtaining the alternative background induced systematic uncertainties for this best-fitting model, we adopt the procedure of Acero et al. (2016b). The authors of this study have developed a method to calculate these uncertainties, in which they employ alternative interstellar emission models in addition to their standard (STD) model. Using the notation of Ackermann et al. (2018), the systematic error is then calculated as

$$\delta P_{\rm sys} = \sqrt{\frac{1}{\sum_i \sigma_i^{-2}} \sum_i \sigma_i^{-2} (P_{\rm STD} - P_i)^2}. \quad (6)$$

In this equation, $P_{\rm STD}$ is the value for which we want to find the systematic error of, i.e. the flux value when using the standard background model. The flux values and their statistical errors when using the alternative background models are P_i and σ_i, respectively. In our case, we used the *Fermi* background as our standard model and the 24 FEMs as our alternatives. The resulting systematic uncertainties are shown as the green error bars in Fig. 10. For our case the statistical errors dominate over the systematic uncertainties. The reason for these large errors in our case as opposed to the much smaller errors in the recovered spectrum in the original study of the M31 excess (Ackermann et al. 2017a) is the following. In that study, the individual normalizations of the point sources in the bin-by-bin analysis were fixed to their best-fitting values obtained in their broad-band (i.e. over the whole energy range) analysis. We stayed agnostic to the overall shape of the spectrum by only performing a bin-by-bin analysis and keeping the normalizations of all sources in our ROI completely free, since we wanted to understand how this would affect the results. It turns out that the statistics are too low to make a definite statement about this. It is interesting however that the excess survives all parts of our analysis with our rigorous and agnostic treatment of the data.

4.3 Implications for the DM interpretation

We have tested how strong the DM hypothesis is and if it can survive the inclusion of our more physically motivated models. We do this with a template-nesting approach similar to the one used in Di Mauro et al. (2019). Even if we assume that parts of the emission are due to annihilating DM, there is still the emission of astrophysical nature, which has to be accounted for. We therefore successively include templates for emission by MSPs in the disc and bulge of M31, as well as a template for gas- and dust-correlated emission into the

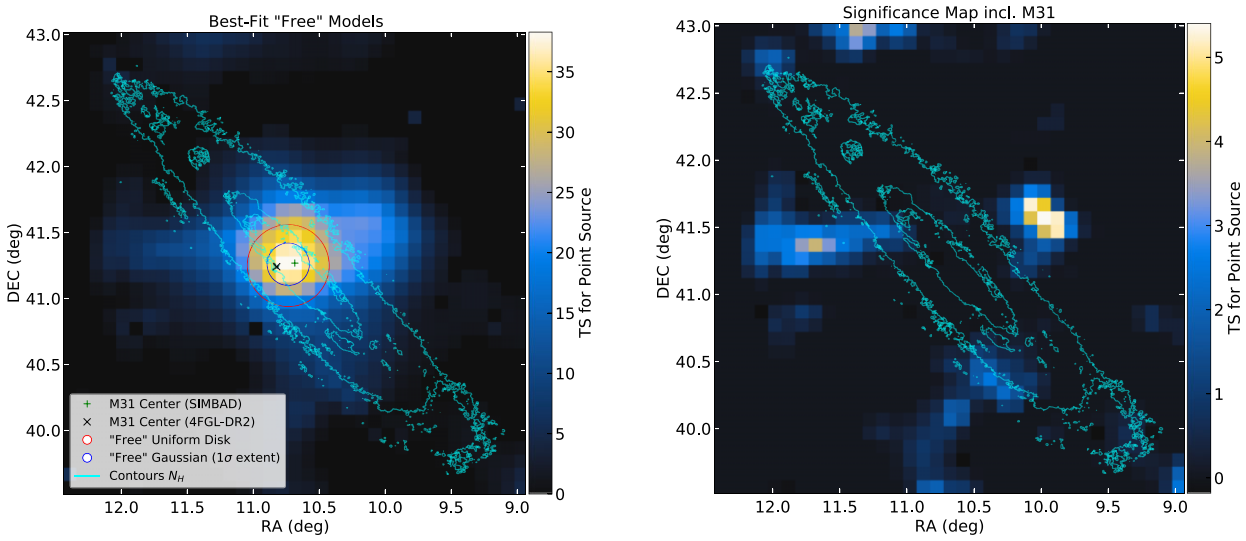

Figure 8. (a) Best-fitting position of the phenomenological models for M31, where we scanned over a grid of spatial locations to find the best-fitting position. (b) The significance map generated by moving a putative point source through a grid of locations when including the source catalogue model for M31. There is no significant unmodelled excess that could be interpreted as a point-like source.

Table 4. Significances for our physically motivated models and the two-component stellar maps. The TS and σ values separated by the slash symbol refer to the Templates separated by the slash symbol in the same order. E.g. the right TS and σ value in the last row corresponds to the Stellar Disc (11–13 Gyr) + W2 Bulge template.

M31 template(s)	TS	σ	$N_{d.o.f.}$
Braun/BraunV2/BraunV3	17.72/15.92/16.63	2.89/2.66/2.75	9
Herschel/Spitzer	18.47/18.95	2.98/3.04	9
150 MHz Radio	9.26	1.68	9
NFW Dark Matter	38.42	5.00	9
W1/W2 Bulge	43.79/43.58	5.44/5.43	9
Stellar Disc (3–5 Gyr)	13.82	2.38	9
Stellar Disc (5–7 Gyr)	14.16	2.42	9
Stellar Disc (7–9 Gyr)	13.69	2.36	9
Stellar Disc (9–11 Gyr)	13.02	2.26	9
Stellar Disc (11–13 Gyr)	12.65	2.21	9
Stellar Disc (3–5 Gyr) + W1/W2 Bulge	46.38/45.92	4.75/4.72	2 × 9
Stellar Disc (5–7 Gyr) + W1/W2 Bulge	46.28/45.81	4.75/4.71	2 × 9
Stellar Disc (7–9 Gyr) + W1/W2 Bulge	46.30/45.88	4.75/4.72	2 × 9
Stellar Disc (9–11 Gyr) + W1/W2 Bulge	40.32/45.39	4.25/4.67	2 × 9
Stellar Disc (11–13 Gyr) + W1/W2 Bulge	45.78/45.36	4.71/4.67	2 × 9

FEM. The chosen templates are based on the significance ranking established in Section 4.1.2 and summarized in Table 4.

The resulting changes in the TS value of the DM hypothesis can be seen in Fig. 11. If we only use the *Fermi* FEM and no additional components are included, the TS value reaches a $\sim 5\sigma$ significance, but quickly falls off significantly ($\sim 1\sigma$) as we account for the astrophysical emission from putative MSPs in the bulge and disc of M31 as well as emission from gas and dust. The purpose of this test was to see whether there was significant emission left for the DM template to 'soak up' after all other anticipated astrophysical emission was covered. The inversion of this test – i.e. including the DM template into the FEM – is of little interest to us, as this would place the priority of the DM component above the astrophysical emission and would make the interpretation of the resulting significances difficult.

In summary, there is no support for DM-related emission once the stellar templates are included.

4.4 Luminosity

Using the spectrum of the best-fitting model, we calculate the γ-ray luminosity of M31 as follows. We calculate the flux for each bin by assuming a power law anchored to $E_0 = 1000\,\mathrm{MeV}$ as

$$F_{\mathrm{bin_i}} = N_0 \int_{E_{\mathrm{bin_i,min}}}^{E_{\mathrm{bin_i,max}}} \frac{\mathrm{d}N}{\mathrm{d}E}\mathrm{d}E = \frac{L}{4\pi d_{\mathrm{M31}}^2} \int_{E_{\mathrm{bin_i,min}}}^{E_{\mathrm{bin_i,max}}} \left(\frac{E}{E_0}\right)^{-\gamma} \mathrm{d}E, \quad (7)$$

where we have expressed the normalization constant N_0 in terms of the luminosity L, the distance to M31 d_{M31}, and the photon energy flux above 500 MeV G_{500}, which is obtained with

$$G_{500} = \int_{E_{\mathrm{min}}}^{E_{\mathrm{max}}} E \left(\frac{E}{E_0}\right)^{-\gamma} \mathrm{d}E, \quad (8)$$

with the limits of the whole energy range considered in this work of $E_{\mathrm{min}} = 500\,\mathrm{MeV}$ and $E_{\mathrm{max}} = 10^5\,\mathrm{MeV}$. We can then optimize the χ^2 quantity

$$\chi^2 = \sum_i \frac{(F_{\mathrm{bin_i}} - F_{\mathrm{bin_i,mod}})^2}{F_{\mathrm{bin_i,err}}^2}, \quad (9)$$

where $F_{\mathrm{bin_i,mod}}$ are the flux values of the model for each bin and $F_{\mathrm{bin_i,err}}$ its associated statistical errors. The optimization was performed with the MINUIT algorithm (Nelder & Mead 1965), available in the IMINUIT package.[10] With this we obtain a total luminosity for

[10] https://iminuit.readthedocs.io/en/stable/

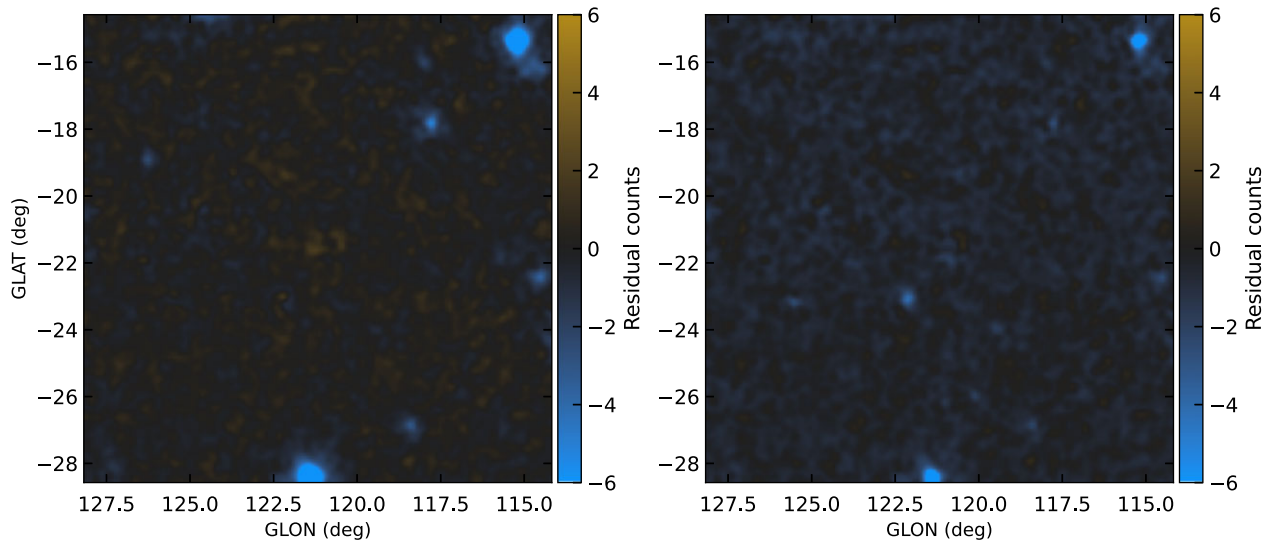

Figure 9. Residual maps when we only fit the *Fermi* foreground model (left-hand panel) to the data and if we include our best-fitting bulge+disc model (right-hand panel). The maps have been smoothed with a Gaussian kernel covering 1°, approximately corresponding to the PSF of *Fermi*-LAT at lower energies.

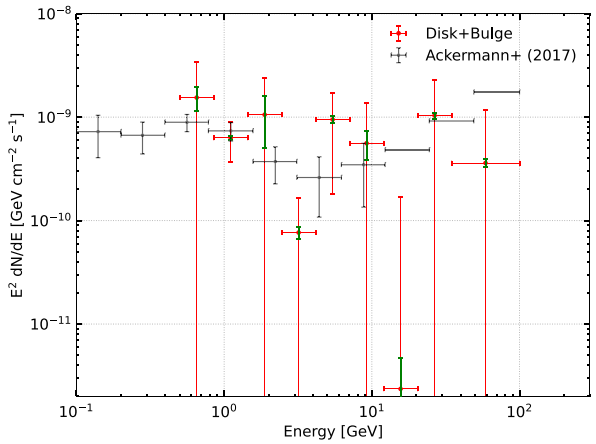

Figure 10. Spectrum of the fit with M31 modelled with the stellar disc (3–5 Gyr) + W1 bulge template, representing a putative MSP population. The red and green error bars represent the statistical and systematical errors, respectively. The flux values when using a uniform disc template of a previous analysis from Ackermann et al. (2017a) are shown in black.

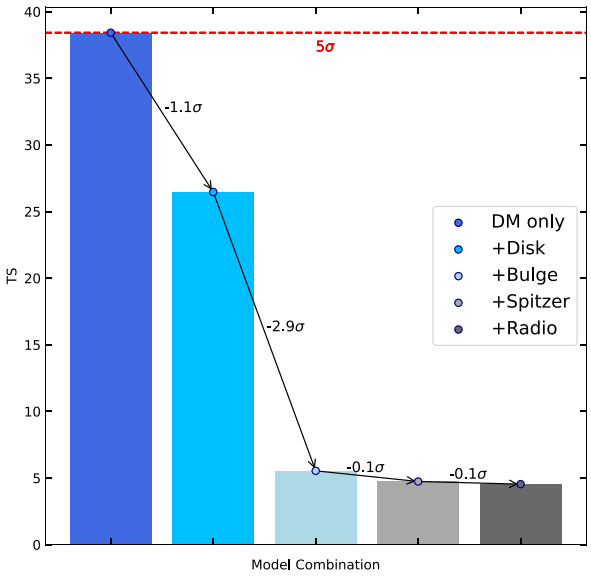

Figure 11. The decrease in significance of detection of the DM template, when we one-by-one include a stellar disc, a bulge, a gas, and a radio template.

M31 of

$$L_{\gamma,\mathrm{M31}} = (1.6 \pm 0.7) \cdot 10^{38}\,\mathrm{erg\,s^{-1}}\,, \tag{10}$$

a luminosity for the stellar bulge of

$$L_{\gamma,\mathrm{M31,B}} = (1.1 \pm 0.2) \cdot 10^{38}\,\mathrm{erg\,s^{-1}}\,, \tag{11}$$

and a 95 per cent C.L. upper limit for the stellar disc of

$$L_{\gamma,\mathrm{M31,D}} < 3.2 \cdot 10^{38}\,\mathrm{erg\,s^{-1}}\,. \tag{12}$$

Note that the disc luminosity is only an upper limit since it was not significantly detected, while, on the other hand, a bulge signal is detected at better than 5σ confidence.

In order to gauge the γ-ray emissivity per unit mass of the M31 stellar structures, we note that the total mass of the galaxy (Tamm et al. 2012) is (1.0–1.5) $\times\ 10^{11}\,\mathrm{M}_\odot$ which we shall understand below as

$$M_{\star,\mathrm{M31}} = (1.25 \pm 0.25) \times 10^{11}\,\mathrm{M}_\odot\,, \tag{13}$$

of which ~ 56 per cent is the disc

$$M_{\star,\mathrm{M31,D}} = (7.0 \pm 1.4) \times 10^{10}\,\mathrm{M}_\odot\,, \tag{14}$$

with the remainder in the bulge (+ halo, using the notation of Tamm et al. 2012). A recent, more tightly constrained determination (Blaña Díaz et al. 2018) for the stellar mass of the M31 bulge (assuming an NFW form for its DM distribution) is that it is

$$M_{\star,\mathrm{M31,B}} = (3.0^{+0.1}_{-0.2}) \times 10^{10}\,\mathrm{M}_\odot\,. \tag{15}$$

From these measurements we infer the following central values (upper limit in the case of the disc) for the γ-ray emissivity per

unit stellar mass

$$L_{\gamma,\star,M31} = (1.3 \pm 0.6) \cdot 10^{27}\,\mathrm{erg\,s^{-1}} \quad (16)$$

$$L_{\gamma,\star,M31,B} = (3.6 \pm 0.3) \cdot 10^{27}\mathrm{erg\,s^{-1}} \quad (17)$$

$$L_{\gamma,\star,M31,D} < 5.2 \cdot 10^{27}\mathrm{erg\,s^{-1}}\,. \quad (18)$$

4.5 MSP interpretation

Here, we consider whether MSPs belonging to M31 can supply its detected γ-ray signal. The stellar population of M31 is predominantly old with its bulge and disc populations of similar mean age (Tamm et al. 2012). On the basis of the data presented by Williams et al. (2017), we determine a mass-weighted mean stellar age of 10–11 Gyr across the entire galaxy. At such an age, the recent binary population synthesis modelling by Gautam et al. (2022) suggests that the MSP population of M31 numbers around 10^6, and the spin-down power liberated by magnetic braking of these pulsars is $\sim 5 \times 10^{28}$ erg s^{-1} M$_\odot$; this is easily sufficient to sustain the observed γ-ray luminosity.[11]

Despite the energetics being easily met, two question marks hang over any putative identification of the M31 γ-rays with an MSP signal: (i) the rather hard spectrum shown in the SED plot (Fig. 10) does not resemble the classical $\sim$2 GeV bump spectrum of γ-ray MSPs, pulsars, or, indeed, the GCE; and (ii) despite the strong detection of the bulge, the disc signal is only marginally detected. Given, as already noted, these stellar populations are characterized by similar mean ages, why does the disc not produce an easily detectable γ-ray signal from its MSPs given its stellar mass actually exceeds that of the bulge?

Of course, we should grant at the outset that, formally, neither of these points is a show-stopper given the scale of present uncertainties/errors in our analysis (i.e. the large error bars on the γ-ray SED we have measured would accommodate a bump-like signal and the 95 per cent upper limit on the disc γ-ray efficiency is clearly compatible with the measured efficiency of the bulge). Nevertheless, in anticipation of tighter constraints, we point out here that there is a natural scenario involving MSPs that would accommodate both a hard SED from the M31 bulge and a small, perhaps undetectable, signal from the disc. This, namely, is that the overall bulge signal is dominated not by the aggregate, 'prompt' curvature radiation from the magnetospheres of all the individual (albeit unresolved) MSPs, but rather by the inverse-Compton upscattering of ambient photons by the cosmic ray electrons and positrons launched by these MSPs into the bulge ISM (cf. Petrović, Serpico & Zaharijas 2015; Ackermann et al. 2017a; Song et al. 2021). Such emission can easily reproduce the observed hard spectrum up to 10s of GeV (e.g. Gautam et al. 2022). A natural explanation, then, of why we do not detect (or only marginally detect) such IC emission from the disc is that, in this environment, the ratio u_B/u_{ISRF} of the magnetic field to the local interstellar radiation field energy densities is significantly smaller than in the bulge (see below). Indeed, consistent with this, the high concentration of stars in the M31 bulge guarantees that the ISRF should reach a peak in this region.

To put some numbers to this, on the basis of the stellar luminosity we (Roth, Krumholz and Crocker, in preparation) calculate a total radiation field energy density in the M31 bulge of 12 and a 0.7 eV cm^{-3} in the disc at the approximate radius of the star-forming ring (10 kpc). For the magnetic field, according to radio continuum studies (and making the equipartition assumption) (Hoernes, Beck & Berkhuijsen 1998; Gießübel & Beck 2014), the field amplitude in M31 reaches a peak at a radius of 800–1000 pc into the bulge; the measured value here, 19 ± 3 μG, thus defines an upper limit to the ISM magnetic field in the bulge region from where the measured γ-ray signal emerges. On the other hand, the magnetic field amplitude at $\sim$10 kpc radius coincident with the star-forming ring is '3–4 times' (Hoernes et al. 1998) smaller than this. Overall, we thus infer

$$\left.\frac{u_B}{u_{\mathrm{ISRF}}}\right|_{\mathrm{M31,B}} < 0.8 \quad , \quad \left.\frac{u_B}{u_{\mathrm{ISRF}}}\right|_{\mathrm{M31,D}} \sim 1.1\,, \quad (19)$$

which implies a transition from IC loss dominance in the bulge to synchrotron loss dominance in the disc, consistent with our scenario (albeit poorly constrained given the uncertainties). An additional, practical consideration is that a postulated γ-ray signal related to the disc stars would be distributed over a significantly larger solid angle than the bulge signal and may be soaked up in one or other of the extended foreground templates. An obvious extension to our analysis here would be to introduce a simultaneous spectro-morphological treatment of the non-thermal radio data covering M31 (e.g. Tabatabaei et al. 2013) in addition to the γ-ray data, which is beyond the scope of this paper.

Overall, our results are thus consistent with previous findings that the M31 signal, despite the hard spectrum (assuming the central values of the data points are close to correct) and despite the weak or even (consistent with) null detection of the massive stellar disc, is completely consistent with a dominantly MSP origin.

5 DISCUSSION AND CONCLUSIONS

In this work, we have performed a reanalysis of 10 yr of *Fermi*-LAT data in the GeV energy range from the M31 region. We have employed novel image reconstruction techniques to derive a multitude of alternative foreground models, giving us a robust framework for the evaluation of the systematic uncertainties in the Galactic diffuse gamma-ray emission models. We have concentrated on the morphology of the emission of M31 by constructing empirical templates, which best represent the stellar population in the bulge and disc of M31. We have opted for this approach instead of using analytical functions like the Einasto profile to avoid possible morphological degeneracies with the DM distribution – as it is popularly modelled by density distributions similar to Einasto profiles.

The main results from our analyses, which involved the use of multiple foreground models, can be summarized as follows:

(i) In this work, we used *Fermi*-LAT photon events with improved directional reconstruction at low energies (PSF3 for the first three energy bins). One of our main goals was to have a better handle on the spatial morphology of the signal, whereas at low energies the angular resolution of the LAT instrument worsens. This implied a trade-off between angular resolution and statistical power of the recovered signal for M31. Additionally, we varied the fluxes of all sources included in our ROI. These two factors combined implied larger (statistical) error bars for the energy spectrum of M31, compared to previous studies. Based on this spectrum we could determine a γ-ray

[11] Note that Gautam et al. (2022) model a specific channel for the production of MSPs, viz. the accretion-induced collapse (AIC) of O-Ne white dwarfs. Were other channels to the production of MSPs in operation in this old stellar population (e.g. 'recycling' of old, core-collapse neutron stars due to accretion from a binary companion), the spin-down power available would be even larger. AIC is, however, perhaps singularly good at delivering new MSPs at time separations of many Gyr post star formation; see (Ruiter et al. 2019).

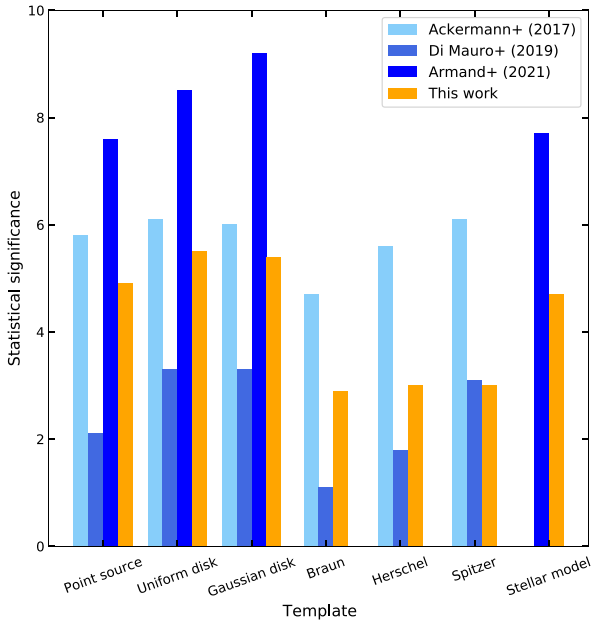

Figure 12. Overview of our results compared to those presented in previous literature dealing with models comparable to those considered in our work. The TS values from the other displayed studies have been converted to statistical significances (in units of σ) via the framework described in Appendix A.

luminosity of $L_{\gamma,\mathrm{M31}} = (1.6 \pm 0.7) \cdot 10^{38}\,\mathrm{erg\,s^{-1}}$, which lies within the range of values compatible with the results in e.g. Ackermann et al. (2017a).

(ii) The M31 excess was robustly detected despite the large set of alternative foreground models used in our analysis. By testing models for the signal (e.g. stellar templates) with different approaches of modelling the γ-ray foreground emission, we could estimate the systematic uncertainties related to the modelling of these complex diffuse emissions. We conclude that it is unlikely that the M31 excess is caused by statistical fluctuations of the Galactic foreground. We also found that that this excess emission appears localized within the bulge and disc regions of M31.

(iii) We have tested stellar templates specifically constructed for this region of the sky to model the emission of a putative population of MSPs in the bulge and disc of M31. We have found these to be as strong as the widely used phenomenological models – a circular disc template with a uniform or Gaussian brightness profile – both sitting at the 5.4σ level. With future observations, we expect that more accurate stellar templates will be able to outperform the more simple phenomenological models.

(iv) Our findings do not support the star formation scenario, in which the regions rich in gas and dust contribute to the gamma-ray emission, since we do not detect the Spitzer and Herschel templates. As we also do not detect any of the hydrogen maps (Braun, BraunV2 and BraunV3) individually, the scenario where the main contribution comes from hadronic-only processes is unlikely as well. These findings are consistent with previous works (e.g. Ackermann et al. 2017a) and can be linked to the properties of M31, specifically to the fact that the SFR of M31 was found to be about 10 times smaller compared to that of the MW (Ford et al. 2013), which can lead to a decreased contribution to the emission from the disc.

(v) We compare all the models considered in this work with relevant previous literature, where possible, in Fig. 12. We can see that our work clearly separates the phenomenological and stellar models from the gas and dust templates. Indeed, the significances we derive for the Point source, uniform, and Gaussian disc results are comparable to those of Ackermann et al. (2017a), while the Braun, Herschel, and Spitzer template results are comparable to those found by Di Mauro et al. (2019). In older analyses and treatment of the data, this distinction between the phenomenological models and the gas and dust templates was not as significant.

(vi) We have tested whether the DM hypothesis survives the inclusion of the stellar templates. By nesting the stellar and DM templates during the fit, we find that the significance of the DM template drops to the 1σ level. We therefore come to a similar conclusion as in related studies using this methodology [e.g. in Di Mauro et al. (2019), Armand & Calore (2021) for M31 and e.g. in Macias et al. (2019) for the MW or others (e.g. Beck & Colafrancesco (2017)]: There is no significant leftover γ-ray emission which can be attributed to annihilating DM in the centre of M31, when maps of stellar mass are included in the fit. However, the issue of degeneracy between the DM and bulge template is still present. In this work, we do not address this part specifically, but instead, focus on showing the unsolicited nature of the DM template, if the well-motivated and observationally supported astrophysical emission is accounted for.

(vii) We used the results of the binary population synthesis modelling by Gautam et al. (2022) to determine whether the MSPs theory could explain the observed luminosity per stellar mass, hard spectrum, and high stellar bulge-to-disc flux ratio. We found that (i) the observed energetics are easily met if M31 hosts about 10^6 unresolved pulsars with an average spin-down luminosity of $\sim 5 \times 10^{28}\,\mathrm{erg\,s^{-1}\,M_\odot}$, and (ii) since the electrons injected by the MSPs lose energy more efficiently in the bulge than in the disc through IC, it is reasonable to conclude that both the hard spectrum and high bulge-to-disc ratio could be explained by an MSPs IC emission scenario. A potential detection of a high-energy tail (at TeV energies) in the M31 spectrum would provide strong support for such a scenario. In future work, we will perform a sensitivity analysis similar to those by e.g. Macias et al. (2021) and Song, Macias & Horiuchi (2019) to investigate the capabilities of the upcoming Cherenkov Telescope Array to a tentative high-energy tail from MSPs in M31.

ACKNOWLEDGEMENTS

We would like to thank Giuseppe Puglisi for all the help he provided with his inpainting package and Nicolas Martin for providing results from his previous works. We would also like to thank Ariane Dekker, Ebo Peerbooms, Sahaja Kanuri, and Manuel Loparco for fruitful discussions and their support. OM was supported by the GRAPPA Prize Fellowship. The work of SA was supported by JSPS/MEXT KAKENHI Grant Numbers JP17H04836, JP20H05850, and JP20H05861. The work of SH is supported by the U.S. Department of Energy under the award number DE-SC0020262, NSF Grant numbers AST-1908960 and PHY-1914409, and JSPS KAKENHI Grant Number JP22K03630. This work was supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abazajian K. N., Kaplinghat M., 2012, Phys. Rev. D, 86, 083511
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., 2016a, ApJS, 223, 26
Acero F. et al., 2016b, ApJS, 224, 8
Ackermann M. et al., 2015, ApJ, 799, 86
Ackermann M. et al., 2017a, ApJ, 836, 208
Ackermann M. et al., 2017b, ApJ, 840, 43
Ackermann M. et al., 2018, ApJS, 237, 32
Ajello M. et al., 2016, ApJ, 819, 44
Armand C., Calore F., 2021, Phys. Rev. D, 103, 083023
Ballet J., Burnett T. H., Digel S. W., Lott B., 2020, preprint (arXiv:2005.11208)
Bartels R., Storm E., Weniger C., Calore F., 2018, Nat. Astron., 2, 819
Beck G., Colafrancesco S., 2017, J. Cosmol. Astropart. Phys., 10, 007
Bertone G., Hooper D., Silk J., 2005, Phys. Rep., 405, 279
Blaña Díaz M. et al., 2018, MNRAS, 481, 3210
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Braun R., Thilker D. A., Walterbos R. A. M., Corbelli E., 2009, ApJ, 695, 937
Bucher M., Racine B., van Tent B., 2016, J. Cosmol. Astropart. Phys., 05, 055
Burns A.-K., Fieg M., Karwin C. M., Rajaraman A., 2021, Phys. Rev. D, 103, 063023
Calore F., Cholis I., McCabe C., Weniger C., 2015, Phys. Rev. D, 91, 063003
Chan M. H., Yeung C. F., Cui L., Leung C. S., 2021, MNRAS, 501, 5692
Charles E. et al., 2016, Phys. Rept., 636, 1
Chernoff H., 1954, Ann. Math. Stat., 25, 573
Coleman B., Paterson D., Gordon C., Macias O., Ploeg H., 2020, MNRAS, 495, 3350
Conn A. R. et al., 2011, ApJ, 740, 69
Conn A. R. et al., 2012, ApJ, 758, 11
Courteau S., Widrow L. M., McDonald M., Guhathakurta P., Gilbert K. M., Zhu Y., Beaton R. L., Majewski S. R., 2011, ApJ, 739, 20
Daylan T., Finkbeiner D. P., Hooper D., Linden T., Portillo S. K. N., Rodd N. L., Slatyer T. R., 2016, Phys. Dark Universe, 12, 1
Diaz J. D., Koposov S. E., Irwin M., Belokurov V., Evans N. W., 2014, MNRAS, 443, 1688
Di Mauro M., 2021, Phys. Rev. D, 103, 063029
Di Mauro M., Hou X., Eckner C., Zaharijas G., Charles E., 2019, Phys. Rev. D, 99, 123027
Do A., Duong M., McDaniel A., O'Connor C., Profumo S., Rafael J., Sweeney C., Vera W., 2021, Phys. Rev. D, 104, 123016
Dotter A., Chaboyer B., Jevremovic D., Kostov V., Baron E., Ferguson J. W., 2008, ApJS, 178, 89
Draine B. T. et al., 2014, ApJ, 780, 172
Eckner C. et al., 2018, ApJ, 862, 79
Elad M., Starck J.-L., Querre P., Donoho D., 2005, Appl. Comput. Harmon. Anal., 19, 340
Ford G. P. et al., 2013, ApJ, 769, 55
Fragione G., Antonini F., Gnedin O. Y., 2019, ApJ, 871, L8
Gautam A., Crocker R. M., Ferrario L., Ruiter A. J., Ploeg H., Gordon C., Macias O., 2022, Nat. Astron., 6, 703
Gießübel R., Beck R., 2014, A&A, 571, A61
Gordon C., Macias O., 2013, Phys. Rev. D, 88, 083521
HI4PI Collaboration, 2016, A&A, 594, A116
Hoernes P., Beck R., Berkhuijsen E. M., 1998, in Sofue Y, ed., Proc. IAU Symp. 184, Properties of synchrotron emission and magnetic fields in the central region of M31. Kluwer, Dordrecht, p. 351
Hooper D., Goodenough L., 2011, Phys. Lett., B697, 412
Ibata R. A. et al., 2014, ApJ, 780, 128
Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Poppel W. G. L., 2005, A&A, 440, 775
Karwin C., Murgia S., Campbell S., Moskalenko I., 2019, Proc. Sci., 36th International Cosmic Ray Conference. SISSA, Trieste, PoS#570
Karwin C. M., Murgia S., Moskalenko I., Fillingham S., Burns A.-K., Fieg M., 2021, Phys. Rev. D, 103, 023027
Kormendy J., 2013, Secular Evolution in Disk Galaxies. Cambridge Univ. Press, Cambridge
Li H., 2014, Smile. Available at https://haifengl.github.io
Licquia T. C., Newman J. A., 2015, ApJ, 806, 96
Lisanti M., Mishra-Sharma S., Rodd N. L., Safdi B. R., 2018, Phys. Rev. Lett., 120, 101101
Lockman F. J., Free N. L., Shields J. C., 2012, AJ, 144, 52
Lorimer D. R. et al., 2006, MNRAS, 372, 777
Macias O., Gordon C., 2014, Phys. Rev. D, 89, 063515
Macias O., Gordon C., Crocker R. M., Coleman B., Paterson D., Horiuchi S., Pohl M., 2018, Nature 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
Martin N. F., Ibata R. A., McConnachie A. W., Mackey A. D., Ferguson A. M. N., Irwin M. J., Lewis G. F., Fardal M. A., 2013, ApJ, 776, 80
Martin N. F. et al., 2014, ApJ, 787, 19
McConnachie A. W. et al., 2018, ApJ, 868, 55
McDaniel A., Jeltema T., Profumo S., 2019, Phys. Rev. D, 100, 023014
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Nelder J. A., Mead R., 1965, Comput. J., 7, 308
Ness M., Lang D., 2016, AJ, 152, 14
Petrović J., Serpico P. D., Zaharijas G., 2015, J. Cosmol. Astropart. Phys., 2015, 023
Pohl M., Macias O., Coleman P., Gordon C., 2022, ApJ, 929, 136
Puglisi G., Bai X., 2020, ApJ, 905, 143
Recchia S., Gabici S., Aharonian F. A., Niro V., 2021, ApJ, 914, 135
Ruiter A. J., Ferrario L., Belczynski K., Seitenzahl I. R., Crocker R. M., Karakas A. I., 2019, MNRAS, 484, 698
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Self S. G., Liang K.-Y., 1987, J. Am. Stat. Assoc., 82, 605
Song D., Macias O., Horiuchi S., 2019, Phys. Rev. D, 99, 123020
Song D., Macias O., Horiuchi S., Crocker R. M., Nataf D. M., 2021, MNRAS, 507, 5161
Storm E., Weniger C., Calore F., 2017, J. Cosmol. Astropart. Phys., 08, 022
Strong A. W., Moskalenko I. V., 1998, ApJ, 509, 212
Sudoh T., Linden T., Beacom J. F., 2021, Phys. Rev. D, 103, 083017
Tabatabaei F. S., Berkhuijsen E. M., Frick P., Beck R., Schinnerer E., 2013, A&A, 557, A129
Tamm A., Tempel E., Tenjes P., Tihhonova O., Tuvikene T., 2012, A&A, 546, A4
Ulyanov D., Vedaldi A., Lempitsky V., 2020, Int. J. Comput. Vis., 128, 1867
Watkins L. L., Evans N. W., An J. H., 2010, MNRAS, 406, 264
Wilks S. S., 1938, Annals Math. Statist., 9, 60
Williams B. F. et al., 2017, ApJ, 846, 145
Wright E. L. et al., 2010, AJ, 140, 1868

APPENDIX A: STATISTICAL FRAMEWORK

To test for the significance of the model under consideration, we used a bin-by-bin analysis, where we originally divided the data into 10 logarithmically spaced energy bins, but later combined the last two into one to compensate for the lack of statistical power in these bins, which comes from the low photon count in this high-energy range. For each bin we performed an independent maximum-likelihood fit and the test statistic for each bin is calculated as the likelihood ratio

$$TS_{\rm bin} = -2\ln\frac{L_{\rm null}(\hat{\boldsymbol{\theta}}_{\rm BKG})}{L_{\rm alt}(\hat{\boldsymbol{\theta}}_{\rm M31}, \hat{\boldsymbol{\theta}}_{\rm BKG})}, \tag{A1}$$

where $L_{\rm null}$ is the sum of the Poisson likelihoods (one for each bin) for the background only and $L_{\rm alt}$ is for the hypothesis, where we included the model for M31 on top of the background with $\hat{\boldsymbol{\theta}}_{\rm M31}$ and $\hat{\boldsymbol{\theta}}_{\rm BKG}$ representing the parameters for the M31 model and background,

respectively. For the background we include all sources from the 4FGL-DR2 catalogue in our 10 deg ROI together with a model for the interstellar diffuse gamma-ray emission and the isotropic gamma-ray background. These background models are either the ones from the *Fermi* collaboration or our FEMs.

In Wilks (1938) it was shown that $TS_{\rm bin}$ is distributed as χ^2_k in the null hypothesis. The subscript k refers to the independent random variables left free to vary, which are normally called degrees of freedom. This theorem (known as Wilk's theorem) holds as long as the parameters can live in the whole parameter space, but since our parameters live in $\mathbb{R}^+$,[12] this theorem no longer holds. In this case we have to use the Chernoff Theorem from Chernoff (1954), where they showed that for one degree of freedom the TS distribution will then be a mixture of χ^2 and the Dirac delta function δ as

$$p(TS) = 2^{-1}(\delta(TS) + \chi^2(TS)). \tag{A2}$$

This equation implies the following: Under the null hypothesis, 50 per cent of the time the evaluated amplitude (our parameter left free to vary in the fit) will be negative, in which case the TS will be assigned through the Delta function, and the other 50 per cent of the time the amplitude is non-negative and the TS distribution will follow from Wilk's theorem. As shown in case 9 of Self & Liang (1987), this holds for any number of degrees of freedom n, such that

$$p(TS) = 2^{-n}\left(\delta(TS) + \sum_{i=0}^{n}\binom{n}{i}\chi^2_i(TS)\right). \tag{A3}$$

As before, the delta function ensures that the values are non-negative, whereas the binomial coefficient $\binom{n}{i}$ accounts for all possible combinations of non-negative amplitudes, each following a χ^2 distribution, and the overall factor of 2^{-n} represents all possible arrangements of n energy bins, all having non-negative values. From this, the significance in units of the standard deviation σ can then be calculated as

$$\sigma \equiv \sqrt{\text{InverseCDF}(\chi^2_k, CDF[p(TS), \widehat{TS}])}, \tag{A4}$$

where (Inverse)CDF is the (inverse) cumulative distribution function and the observed TS value is denoted by $\widehat{TS}$, adopting the notation of Macias et al. (2019).

[12]In our analysis, we leave the normalization of the templates free to vary in the fit, but it can not be a negative value.

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