Star-forming galaxies dominate the diffuse, isotropic γ-ray background

Matt A. Roth[1,*], Mark R. Krumholz[1,2], Roland M. Crocker[1], and Silvia Celli[3]

[1]Research School of Astronomy and Astrophysics, The Australian National University, Canberra, Australia

[2]ARC Centre of Excellence for All-Sky Astrophysics in Three Dimensions (ASTRO-3D), Canberra, Australia

[3]Dipartimento di Fisica dell'Università La Sapienza e INFN, P. le Aldo Moro 2, 00185 Rome, Italy

[*]matt.roth@anu.edu.au

ABSTRACT

The *Fermi* Gamma Ray Space Telescope has revealed a diffuse, isotropic γ-ray background at energies ranging from 0.1 GeV to 1 TeV [1] whose astrophysical sources remain uncertain. Previous efforts to understand the origin of this background have been hampered by the lack of physical models capable of predicting the γ-ray emission produced by the many candidate sources, which include star-forming galaxies (SFGs) [2–6], active galactic nuclei (particularly blazars [7–9]), millisecond pulsars [7], and dark matter annihilation [10]. In the absence of predictive models, estimates of the contribution from potential sources have relied on a highly-uncertain process of empirically scaling the emission from a small sample of local, resolved sources by their estimated cosmological abundances. Here we present the first calculation of the contribution of SFGs to the γ-ray background that is based on a physical model for the γ-ray emission produced when cosmic ray ions accelerated in supernova remnants interact with the interstellar medium [11]. We validate the model by showing that it reproduces the γ-ray spectra, source count distribution and far infrared-γ-ray correlation observed for nearby, resolved SFGs. When we apply the model to the observed cosmological SFG population, we recover an excellent match to the γ-ray background from 1 GeV to 1 TeV. Our result shows that SFGs alone can explain the full diffuse γ-ray background over this energy range, and strongly suggests that emission in excess of our model at energies $<$1 GeV originates from cosmic ray electrons produced in the same galaxies.

1 Introduction

The γ-ray sky has been measured exquisitely by the Large Area Telescope (LAT) on board the *Fermi* Gamma Ray Space Telescope over more than a decade. The observed emission can be decomposed into resolved sources – which include pulsar wind nebulae, millisecond pulsars, supernova remnants, and active galactic nuclei (mostly blazars [12]) – and diffuse emission. The diffuse component can be further separated into Galactic foreground emission and an isotropic, extragalactic background [1]. The former is predominantly produced by the interactions of relativistic ions, referred to as cosmic rays (CRs), with the gas that fills the Galaxy between the stars, the interstellar medium (ISM), but the origin of the latter is much more uncertain.

SFGs (including the Milky Way) produce diffuse γ-ray emission as a result of diffusive shock acceleration of CRs in

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

supernova remnants [13]. This process transfers ∼10% of the supernova mechanical energy to relativistic particles, yielding on average $\sim 10^{50}$ erg in CR ions per supernova [14, 15], with a substantially smaller portion deposited in CR electrons. The resulting CRs follow a power law distribution in particle momentum, well approximated for relativistic CRs as a power law in total particle energy E [16, 17], $dN/dE \propto E^{-p}$, with an index p that observations of individual supernova remnants, analytical models, and numerical simulations all indicate to be in the range $p \approx 2.0-2.6$, with a mean value $p \approx 2.2-2.3$ [18, 19]. Some of these CR ions collide inelastically with ISM nuclei, producing roughly equal numbers of π^0, π^+ and π^- mesons that rapidly decay via the channels $\pi^0 \rightarrow 2\gamma$, $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$, and $\pi^+ \rightarrow \mu^+ + \nu_\mu$. The decay of π^0 particles is responsible for most of the observed Galactic γ-ray foreground, which displays a characteristic spectrum that rises sharply from $\sim 0.1-1$ GeV as a result of the 135 MeV rest mass of the π^0 particle.

2 Galactic emission model

The diffuse γ-ray emission from a SFG depends on three factors: its total star formation rate (which determines its supernova rate and thus the rate at which CRs are injected), the distribution of γ-ray energies produced when individual CRs collide with ISM nuclei (which depends on the parent CR energy E) and the fraction of CRs (again as a function of E) that undergo inelastic collisions before escaping the galaxy. The latter is known as the calorimetry fraction $f_{\rm cal}$, and is the most poorly-understood of the three.

For the purposes of the calculations we present here, we compute the conversion from star formation rate to CR injection by assuming that stars form with a Chabrier initial mass function [20], which gives the distribution of mass for newly-formed stars, and that stars with initial masses of $8-50$ $M_\odot$, where $M_\odot$ is the mass of the Sun, end their lives as supernovae [21]. Each supernova injects 10^{50} erg of energy in CR ions, distributed in energy for CR energies $E > m_p c^2$ as $d\dot{N}/dE = \phi\ \dot{M}_*\ (E/E_0)^{-p}\ \exp(-E/E_{\rm cut})$, where $E_0 = 1$ GeV, the cutoff energy $E_{\rm cut} = 10^8$ GeV[1], $\dot{M}_*$ is the galaxy's star formation rate, and the spectral index $p = 2.2$ [18, 19]. The normalisation factor ϕ that corresponds to our choice of initial mass function, supernova mass range, and CR energy per supernova is $\phi = 6.52 \times 10^{42}$ s^{-1} GeV^{-1} $M_\odot^{-1}$ yr.

Similarly, we calculate the γ-ray spectrum produced by CRs of energy E using the parameterised model of Ref. [22], which gives the differential cross section $d\sigma_\gamma/dE_\gamma\left(E_\gamma, E\right)$ for the production of γ-rays of energy E_γ by CRs of energy E in inelastic collisions. Combining these factors, the rate at which a SFG produces γ-ray photons is

$$\frac{d\dot{N}_\gamma}{dE_\gamma}\left(E_\gamma\right) = \int_{m_{\rm p}c^2}^{\infty} \frac{d\sigma_\gamma}{dE_\gamma}\left(E_\gamma, E\right)\ \frac{f_{\rm cal}\left(E\right)}{\sigma_{\rm pp}}\ \frac{d\dot{N}}{dE}\left(E\right)\ dE \tag{1}$$

where $\sigma_{\rm pp} = 40$ mbarn is the mean proton-proton inelastic cross-section. The calorimetry fraction $f_{\rm cal}\left(E\right)$ depends on the properties of the galaxy, and the lack of a model for this dependence has previously precluded proper evaluation of Equation 1. However, Ref. [11] recently introduced a model for $f_{\rm cal}(E)$, based on the idea that the rate at which CRs diffuse through the

[1] The exact choice of the cutoff energy above $E \sim$ PeV makes no practical difference because the injection spectral index $p > 2$, so only a small fraction of the total CR energy is injected at $\gtrsim$ PeV energies regardless of $E_{\rm cut}$, and any CRs that are injected at such high energies produce photons that we do not observe due to $\gamma\gamma$ opacity (see below).

magnetised ISM, and thus the fraction that escape before colliding, is determined by the balance between driving of turbulence by the CR streaming instability and ion-neutral damping. This model successfully reproduces the γ-ray spectra of NGC 253, M82, and Arp 220, and we therefore adopt it for this work. We describe our implementation of the model in detail in the Methods and illustrate the behaviour of $f_{\rm cal}(E)$ over a range of gas surface densities and redshifts in Extended Data Figures 1 and 2.

The γ-rays produced inside a galaxy do not propagate unhindered, but are subject to losses due to pair production in collisions with far-infrared (FIR) photons inside the source galaxy and with photons that are part of the extragalactic background light (EBL) that pervades intergalactic space; the latter is generally the more important effect.

We compute the optical depth $\tau_{\gamma\gamma}$ due to galactic FIR photons using the model of Ref. [23], and the optical depth $\tau_{\rm EBL}$ due to the EBL using the model of Ref. [24]. Taking these into account, Equation 1 can be used to compute the specific photon flux dF_γ/dE_γ (i.e., the number of photons per unit area, time and energy) received at the Earth from a galaxy at redshift z as

$$\frac{dF_\gamma}{dE_\gamma}\left(E_\gamma,z\right) = \frac{(1+z)^2}{4\pi\, d_{\rm L}^2(z)}\,\frac{d\dot{N}_\gamma}{dE_\gamma}\left(E_\gamma(1+z)\right)\; e^{-\tau_{\rm EBL}\left(E_\gamma,z\right)}\; e^{-\tau_{\gamma\gamma}\left(E_\gamma(1+z)\right)} \tag{2}$$

where $d_{\rm L}(z)$ is the luminosity distance of the source. The radiation that is absorbed by the galactic and extragalactic photon fields is reprocessed to lower energies in the pair-production cascade. In this process, both the electron and positron, from the initial high energy pair produced, inverse Compton scatter lower energy photons up to γ-ray energies and these, in turn, produce further pairs, and so on. The photon spectrum from this cascade can be parameterised according to the method developed in Ref. [25]. For the purposes of our calculation here, we include the effect of the cascade by adding a component to dF_γ/dE_γ with a spectral shape as computed by Ref. [25], and with a normalisation such that its energy is equal to the integrated energy lost to photon-photon scattering. In our model the cascade contributes 5 - 20% of the total emission between 1-100 GeV.

3 γ-rays from star-forming galaxies

We now have a model that predicts the γ-ray emission of a SFG. The next step in our analysis is to apply this model to a galaxy survey that samples the SFG population out to the epoch of peak cosmological star formation at $z\sim 2$. For this purpose, we make use of the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) [26] in the GOODS-S field; for each CANDELS galaxy we extract the stellar mass M_*, star formation rate $\dot{M}_*$, and effective radius R_e from the analysis of Ref. [27]. The full sample contains 34,930 galaxies, but we exclude those whose parameters are uncertain because they contain bright active galactic nuclei, have unreliable redshifts, or lack a good fit to the surface brightness profile. This leaves a sample of 22,279 galaxies.

Before we can apply our model for $f_{\rm cal}(E)$ to CANDELS, one further step is required: our model depends on the gas surface density Σ_g, velocity dispersion σ_g, and gas scale height h_g (see Methods), but CANDELS provides only stellar properties, and there is no comparably-large deep survey that provides information on the gas properties. For this reason, we infer $\Sigma_{\rm g}$, $\sigma_{\rm g}$, and h_g from M_*, $\dot{M}_*$, and R_e using observed correlations between galaxies' gas and stellar properties. To convert between

total masses, surface densities, and volume densities, we assume that the galactic disc has an area of $2\pi R_e^2$ and a thickness of $2h_g$. We then derive the gas surface density from the stellar and star formation rate surface densities, $\Sigma_* = M_*/2\pi R_e^2$ and $\dot{\Sigma}_* = \dot{M}_*/2\pi R_e^2$, using the empirical Kennicutt-Schmidt relation obtained by Ref. [28]. Likewise, we derive the gas velocity dispersion from the star formation rate using the empirical relation obtained by Ref. [29]. Finally, we derive the scale height assuming the galactic disc is in hydrostatic equilibrium [30]. We provide full details of these procedures in the Methods.

To verify that this approach predicts reasonably accurate γ-ray spectra, we apply it to four local, resolved galaxies, chosen to span a wide range of gas and star formation surface densities: Arp 220, NGC 253, M31, and NGC 4945. The input data we use for these calculations are summarised in Extended Data Table 1, and we show the results of the computation in Figure 1, where the solid lines show the spectra derived using only stellar data and, for comparison, the dotted lines show the results we obtain if we add directly-measured gas properties. We see that the fits are slightly improved if we make direct use of gas data but, even for the stellar data only, our model reproduces the observed γ-ray spectra to better than a factor of 2 for all galaxies at energies > 1 GeV, and within a factor of ≈ 1.5 for the two more rapidly star-forming galaxies, which, as we show below, are more akin to the population that dominates the γ-ray background. We do not expect agreement at energies $\lesssim 1$ GeV because we have not included the leptonic component of the γ-ray emission, which dominates below the $\sim 0.1-1$ GeV cutoff in the hadronic component set by the π^0 rest mass.

Having verified that we can obtain accurate predictions of γ-ray spectra from stellar data alone, we carry out two additional validation steps. First, we examine the correlation between galaxies' FIR and γ-ray luminosities. We calculate the total γ-ray luminosity in the band $0.1-100$ GeV for each galaxy in our sample by multiplying the integrand in Equation 1 by E_γ and integrating over the γ-ray spectrum. We also convert the star formation rate to an FIR luminosity in the $8-1000$ μm band using the relation in Ref. [3, 31], corrected to a Chabrier IMF [32]. In Figure 2 we show the resulting distribution of galaxies in the $L_{\rm FIR}$-L_γ-plane, along with a power law fit to the data (blue dotted line), $L_\gamma/{\rm erg\ s^{-1}} = 10^{27.69}(L_{\rm FIR}/L_\odot)^{1.19}$. For comparison, the solid green line shows a fit to the 14 nearby SFGs detected in γ-rays by *Fermi* LAT [33]. Our model prediction shows excellent agreement with the observed relation, and we note that both the model and the observed correlation differ noticeably from the calorimetric limit obtained by simply setting $f_{\rm cal}(E) = 1$ in Equation 2 (red dashed line in the figure). Thus the agreement is non-trivial, and suggests that our model is correctly predicting the variation in galaxies' calorimetry fractions as a function of star formation rate.

Our second validation test is to compare our model with counts of *resolved* SFGs observed by *Fermi* LAT. We cannot use the CANDELS catalogue directly for this purpose, because CANDELS has a narrow field of view that provides very little sampling of galaxies at $z \lesssim 0.1$, whereas these local sources are the only ones *Fermi* LAT can resolve. Instead, we perform this test by employing a Monte Carlo scheme to simulate a nearby galaxy population that follows the observed distribution of star formation rates in the local Universe [34, 35], and where the correlation between galaxy star formation rate and γ-ray luminosity is the same as what our model predicts for the low redshift ($z < 1.5$) part of the CANDELS sample. Full details of the Monte Carlo procedure, and the statistical method we use to compute confidence intervals on the observed points, are given

in the Supplementary Information. We show the results of this calculation in Figure 3.

The figure shows that our Monte Carlo simulations predict SFG source counts consistent with observations, with the exception that we do not predict sources as bright as the Milky Way's two satellite galaxies, the Large and Small Magellanic Clouds. This is not surprising, since these are not part of the field galaxy population we are simulating, and are much closer to the Milky Way than we expect any randomly-chosen field galaxy to be.

4 The γ-ray background

Having validated our model, we are now ready to compute the contribution of SFGs to the diffuse, isotropic γ-ray background. We do this by calculating an observed spectrum for each galaxy in the CANDELS sample using Equation 2, and summing over all galaxies in bins of redshift:

$$\Phi\left(E_\gamma\right) = \frac{d^2F_\gamma}{dE_\gamma d\Omega}\left(E_\gamma\right) = \frac{1}{\Omega_S}\sum_{j=1}^{n_{zbin}} f_{corr,j}\sum_{i=1}^{n_{S,j}}\frac{dF_{\gamma,i}}{dE_\gamma}\left(E_\gamma, z_j\right) \tag{3}$$

Here $\Omega_S = 173$ arcmin^2 is the solid angle surveyed by CANDELS, $n_{S,j}$ the number of surveyed galaxies in the jth redshift bin, and n_{zbin} the number of redshift bins. The factor $f_{corr,j}$ is the ratio of the expected total star formation rate in each redshift bin (based on the measured cosmic star formation history [36]) to the sum of the star formation rates of CANDELS galaxies in that bin; its purpose is to correct for the fact that, due to its limited field of view and various observational biases, the distribution of star formation with respect to redshift in CANDELS does not precisely match the total star formation history of the Universe. We use redshift bins of size $\Delta z = 0.1$, chosen to ensure that the number of sample galaxies in each bin is large enough that the uncertainty in the mean spectrum due to Poisson sampling of the galaxy population is small.

We present the results of this calculation in Figure 4. The figure shows that the expected contribution of SFGs to the diffuse isotropic γ-ray background fully reproduces both the intensity and the spectral shape of the observations from ≈ 1 GeV to ≈ 1 TeV. We emphasise that we obtain this agreement from our model with no free parameters: our only inputs are the CR injection spectral index ($p = 2.2$) and the CR energy per supernova (10^{50} erg) that are directly measured in the local Universe, and the distribution of SFGs sampled by CANDELS. The key to the success of the model is the galaxy-by-galaxy calculation of the energy-dependent calorimetry fraction $f_{cal}(E)$. To illustrate this point, we also show (orange dashed line) the result we would obtain simply by setting $f_{cal} = 1$ for all galaxies at all energies, which clearly both overestimates the intensity and yields a spectral slope that is flatter than observed. By contrast, the EBL is important for the spectral shape only for energies > 100 GeV, as illustrated by the green dashed line in Figure 4, which shows the result of a calculation where we omit $\gamma\gamma$ opacity effects.

We show the relative contributions to the background made by galaxies with differing star formation rates and redshifts in Extended Data Figure 3. The figure shows that the background at lower energies is dominated by galaxies from just after cosmic noon ($z \sim 1-2$), while at higher energies where EBL attenuation has a larger effect, the dominant contribution shifts toward lower redshift, so that at 1 TeV the background is dominated by $z \sim 0.1$ sources. At all energies the dominant contribution comes from galaxies at the upper end of the star-forming main sequence, which have high but not extreme star formation rates

for their redshift.

5 Discussion

It is important to put our finding that SFGs dominate the diffuse, isotropic background in the context of recent work, where a number of authors have argued that blazars and other AGN sources contribute significantly or even dominate the background. We find that, while blazars dominate the *resolved* component of the extragalactic γ-ray background, as shown in Figure 3, SFGs dominate the unresolved component. This finding is consistent with statistical analyses of angular fluctuations in the isotropic background and cross-correlations between it and galaxies and quasars, which strongly disfavour blazars as a dominant contributor [4, 7, 37]. Indeed, a straightforward extrapolation of the number counts of observed blazars [38], illustrated by the orange band in Figure 3, also suggests that blazars do not dominate the unresolved background.

Our finding that SFGs alone are able to reproduce the full background also stands in contrast with some earlier analyses (e.g. [3, 39]), though it is consistent with the conclusion of Refs. [5] and [40] that, in the absence of either a physical model for the γ-ray emission of SFGs or a much larger sample of resolved galaxies, it is not possible to rule them out as a dominant contributor. It is nonetheless instructive to examine the precise reasons why our conclusions differ from some earlier work.

One contributing factor is that earlier models were forced to adopt single power laws for the emitted γ-ray spectrum in different classes of galaxies [39, 41]. By contrast, Figure 1 demonstrates that none of the four nearby resolved galaxies shown have spectra that are well described by a γ-ray spectrum in the form of a pure power law over the energy range from $E_\gamma = 1 - 1000$ GeV; our model correctly captures this behaviour, but earlier pure power law models did not. Similarly, we calculate $f_{\rm cal}$ as a function of energy directly, rather than relying on an empirical FIR-γ correlation, and our calorimetry fractions are on average larger than those implicitly assumed in earlier works. This is because many of the lower estimates for the contribution from SFGs to the γ-ray background rely on a FIR-γ relation derived from early *Fermi* detections of < 10 individually-resolved SFGs [3] that yields somewhat lower γ-ray luminosities than more recent fits using a larger (but still small) sample of SFGs [33], and with which our model agrees (Figure 2). Thus the reason we find that SFGs can produce the full background, whereas earlier models could not, is that our model predicts γ-ray emission that is both somewhat brighter and has a more complex spectral shape than the values adopted in earlier work.

Likewise, earlier claims that a variety of other source classes dominate the diffuse isotropic background have also relied on extrapolated empirical correlations with large uncertainties. For instance Ref. [42] estimates the contribution from misaligned active galactic nuclei using a radio-γ relation derived from a sample of 16 resolved objects, coupled with a radio luminosity function extrapolated to redshifts considerably higher than those well-sampled by observations [43]. By contrast, our assignment of γ-ray luminosities to SFGs is based on a physical model that agrees with local observations, and the CANDELS catalogue from which we draw our distribution of SFG properties has very good completeness over the range of redshift and star formation rate that dominates production of the diffuse γ-ray background (see Extended Data Figure 3).

We conclude this discussion with two final questions and one prediction. The first question is whether SFGs can also

explain the background below 1 GeV as arising from the leptonic counterpart to the hadronic emission we have measured. Comparing the total power in the γ-ray background from $0.1-1$ GeV to the total available power in hadronic cosmic rays (the integral under the orange dashed line in Figure 4) indicates that a small fraction of the hadronic power would be sufficient to explain the observed signal, so this explanation is plausible on energetic grounds. However, we leave detailed modelling of the leptonic emission for future work. The second question is whether SFGs also dominate the observed neutrino background at energies > 1 TeV. We address this issue in the Supplementary Information, where we show that our model implies that SFGs can contribute up to $\approx 15\%$ of the neutrino background at 100 TeV if we set $E_{\rm cut} = 100$ PeV, but can be much smaller if $E_{\rm cut} \lesssim 1$ PeV. We show the neutrino spectrum predicted for $E_{\rm cut} = 1$ and 100 PeV in Extended Data Figure 4, but warn that these results are sensitive to both the statistics of nearby, bright sources that are poorly sampled by CANDELS and the precise shape of the CR spectrum at very high energies. We leave a more thorough exploration of this issue for future work as well. Finally, we note that observations by the upcoming Cherenkov Telescope Array [44] and Large High Altitude Air Shower Observatory [45] should significantly extend the population of γ-ray detected SFGs, in addition to probing the spectra of currently detected SFGs to higher energies. Our model makes clear predictions for both source counts and spectral shapes that can be tested against these data.

Supplementary Information is available for this paper. Correspondence and requests for materials should be addressed to the first author. Reprints and permissions information is available at www.nature.com/reprints.

Figures

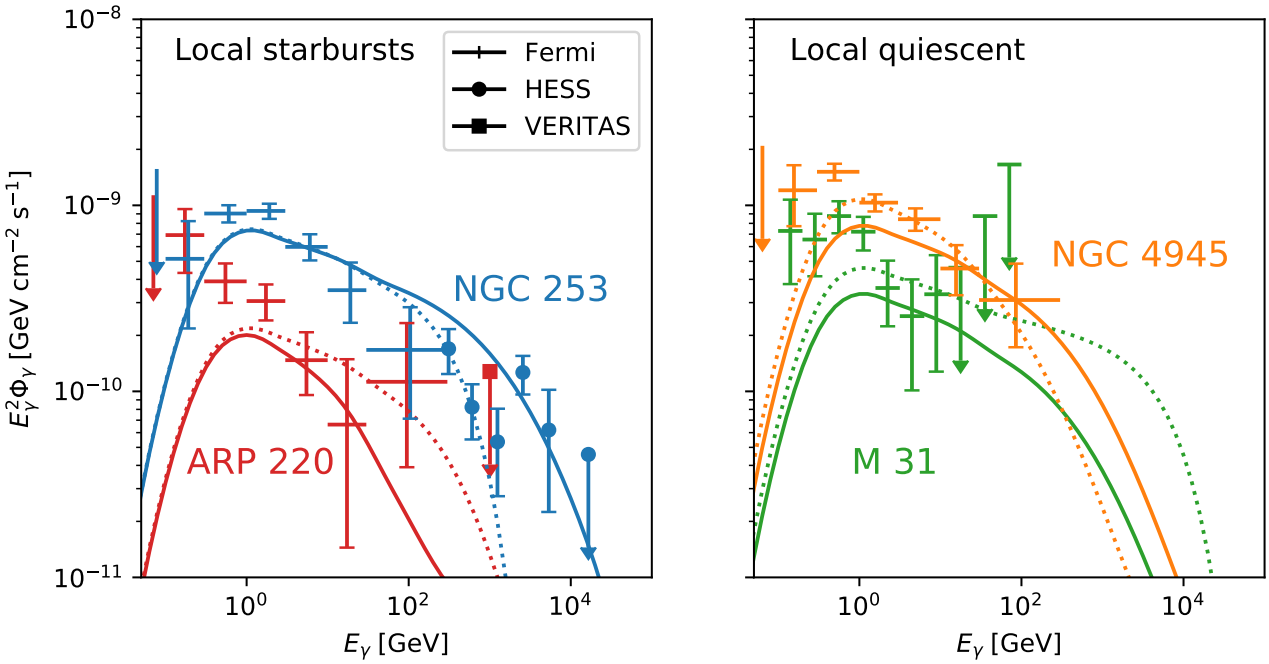

Figure 1. Predicted (lines) and observed (points) γ-ray spectra for a selection of nearby star-forming galaxies detected in γ-rays. The observations shown are taken from a combination of *Fermi* LAT [46], HESS [3], and VERITAS [47]; in the left panel we show the local starburst galaxies Arp 220 and NGC 253, and in the right panel we show the local quiescent galaxies M31 and NGC 4945. The solid lines show model predictions using only stellar data of the type we have available for the CANDELS sample, while the dotted lines shows results predicted if we supplement this with observed gas data. We list the full set of observed quantities used in computing these models in Extended Data Table 1.

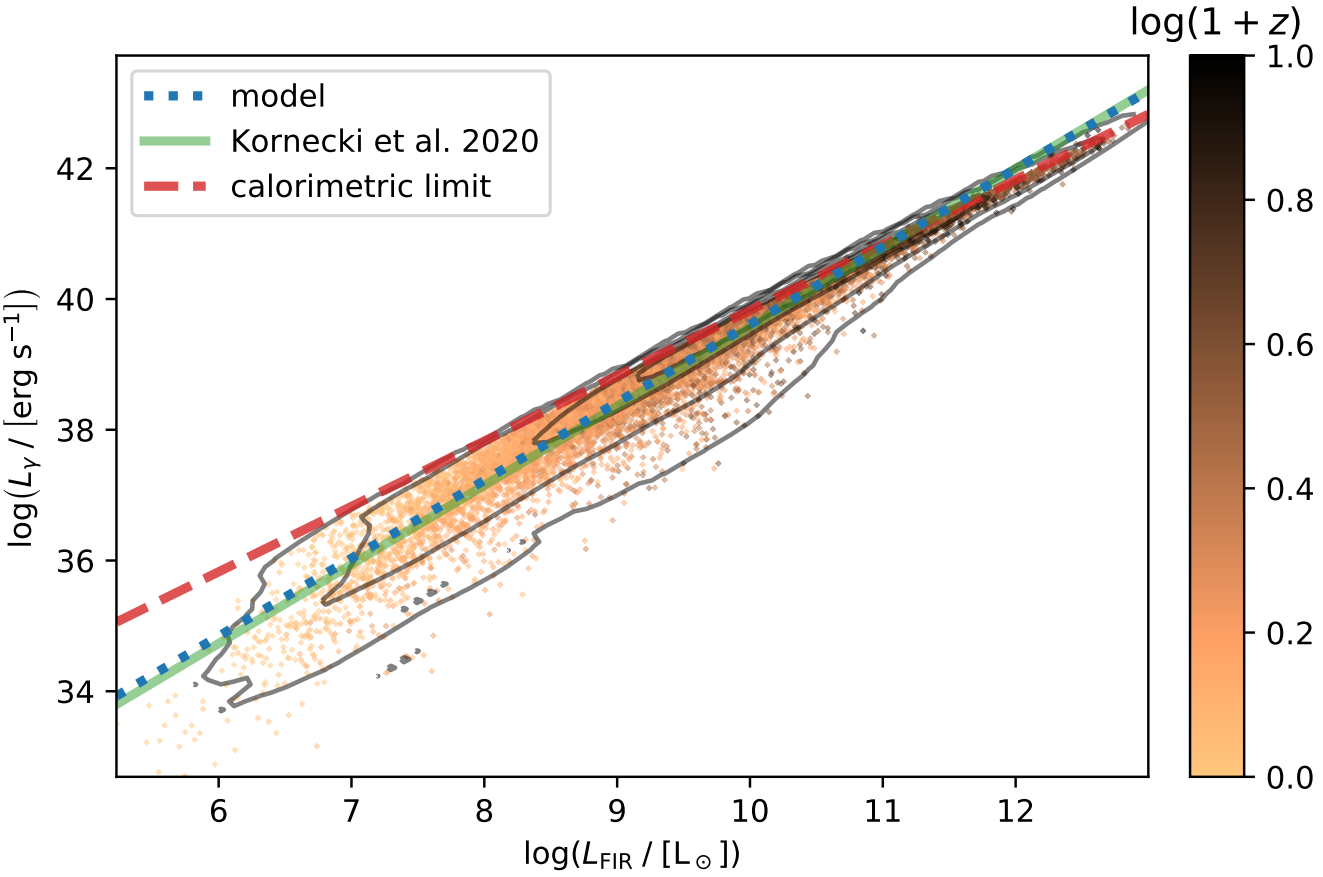

Figure 2. The correlation between far-infrared ($8-1000\ \mu$m) and γ-ray ($0.1-100$ GeV) luminosity for the CANDELS GOODS-S sample, derived using our model. Points show individual CANDELS galaxies, colour-coded by redshift z. Contours show density fractions of 0.002, 0.02, 0.2 and 0.6, computed from a kernel density estimate of the distribution of CANDELS galaxies in the FIR-γ plane. The dashed blue line is a power law fit to the CANDELS sample. For comparison, the solid green line shows the empirical relation measured for 14 nearby, resolved SFGs [38], and the red dashed line is the calorimetric limit obtained by taking $f_{\rm cal}(E) = 1$ at all energies in Equation 1, as obtained by Ref. [3].

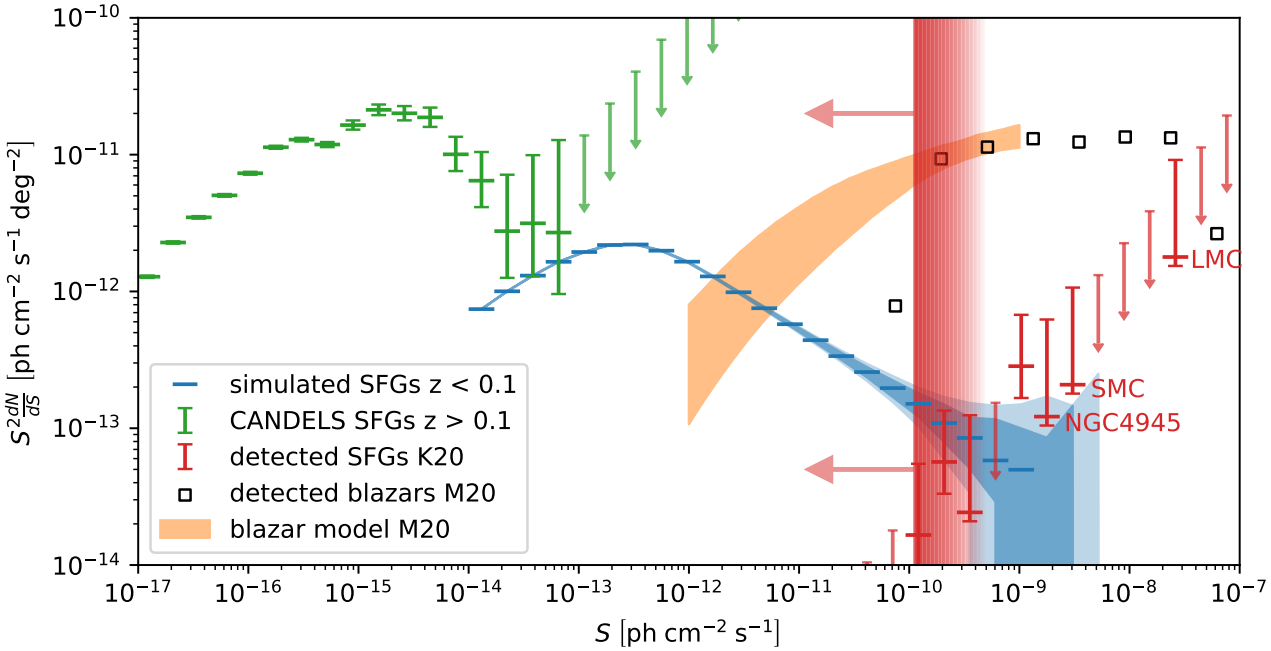

Figure 3. The compensated distribution of γ-ray sources $S^2(dN/dS)$ on the sky as a function of photon flux S integrated from 1 - 100 GeV as seen by *Fermi* LAT. Red and green points show *Fermi*-detected SFGs [33] and model-predicted source counts for observed CANDELS galaxies at $z > 0.1$, respectively; error bars show 90% confidence intervals or upper limits. The three brightest non-empty *Fermi* bins each contain only a single SFG, which we have labelled. Blue points show Monte Carlo realisations of the $z < 0.1$ SFG population, with the light and dark shaded bands indicating 68% and 90% confidence intervals. Black squares show *Fermi*-detected blazars, and the orange band shows the blazar distribution model of Ref. [38]. Finally, the red vertical band indicates the flux range over which *Fermi* observations become incomplete; the left edge of this band is the 4FGL threshold for 98% detection efficiency for sources with spectral index 2.3 [38].

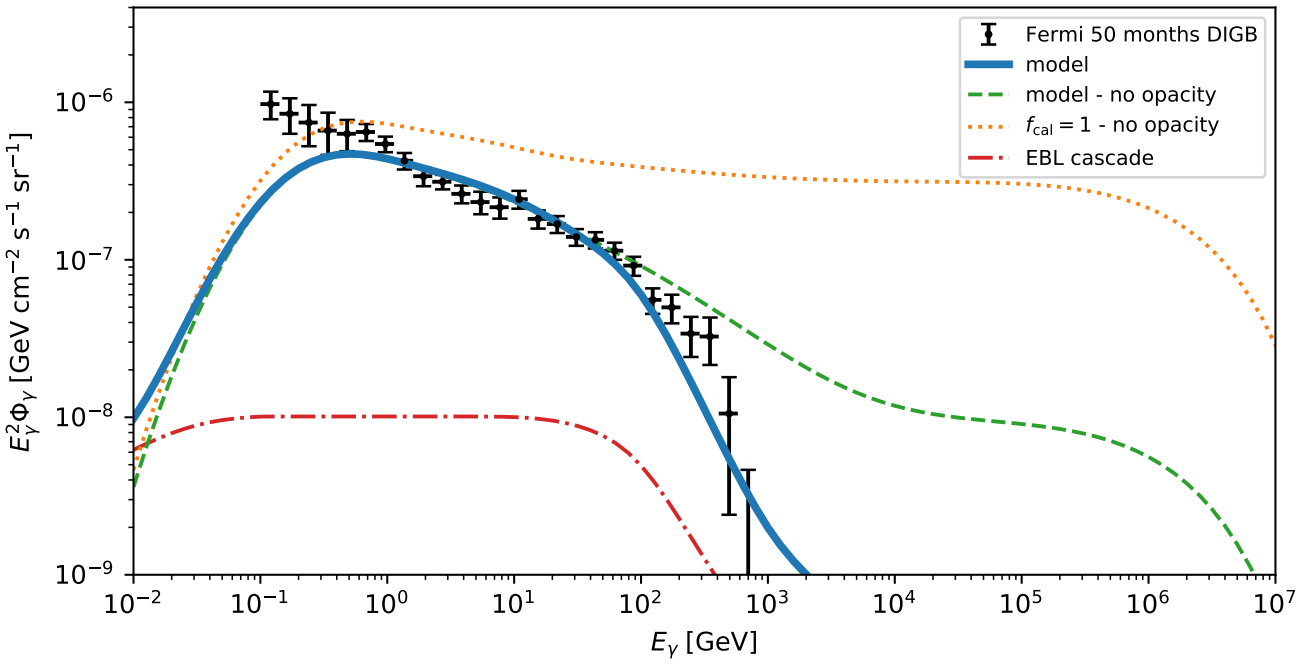

Figure 4. The diffuse isotropic γ-ray background. Black data points show *Fermi* 50-month observations [1]; the blue line shows our prediction for the background due to SFGs; the red dot-dashed line shows the contribution to this spectrum due to the $\gamma\gamma$ scattering cascade, and the green dashed line shows the spectrum we would obtain in the absence of $\gamma\gamma$ opacity. For comparison, the orange line is the model prediction for full calorimetry, i.e., with $f_{\rm cal}(E) = 1$ for all galaxies at all CR energies.

Acknowledgements

This research has made use of the NASA/IPAC Infrared Science Archive, which is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology. Funding for this work was provided by the Australian Government through the Australian Research Council, awards FT180100375 (MRK) and DP190101258 (RMC and MRK), and the Australian National University through a research scholarship (MAR). RMC thanks Oscar Macias and Shin'ichiro Ando for enlightening conversations while a Kavli IPMU-funded guest of the GRAPPA Institute at the University of Amsterdam.

Author contributions

All authors were involved in the design of the study and the interpretation of the results. MAR performed the modelling and data analysis with input from MRK, RMC, and SC. The manuscript was written by MAR, MRK and RMC, and reviewed by all authors.

Methods

Model for calorimetry

We employ the model of Ref. [11] for CR transport and calorimetry in SFGs. The basic premise of the model is that, in the neutral phase that dominates the mass of the ISM and thus the set of available targets for γ-ray production, CR transport is primarily by streaming along magnetic field lines. However, this yields approximately diffusive transport when averaged over scales comparable to or larger than the coherence length of the magnetic field, with a diffusion coefficient $D \approx V_{\rm st} h_g / M_A^3$, where $V_{\rm st}$ is the CR streaming speed, h_g is the gas scale height, and M_A is the Alfvén Mach number of the turbulence. For diffusive transport with losses in a disc geometry, the calorimetry fraction is given by (using the favoured parameters of Ref. [11])

$$f_{\rm cal}(E) = 1 - \left[{}_0F_1\left(\frac{1}{5}, \frac{16}{25}\tau_{\rm eff}\right) + \frac{3\,\tau_{\rm eff}}{4\,M_A^3}\,{}_0F_1\left(\frac{9}{5}, \frac{16}{25}\tau_{\rm eff}\right)\right]^{-1}, \tag{4}$$

where ${}_0F_1$ is the generalised hypergeometric function and $\tau_{\rm eff}$ is the dimensionless effective optical depth of the ISM, given by

$$\tau_{\rm eff} = \frac{\sigma_{\rm pp}\,\eta_{\rm pp}\,\Sigma_g\,h_g\,c}{2\,D_0\,\mu_{\rm p}\,m_{\rm H}}. \tag{5}$$

Here $\sigma_{\rm pp} = 40$ mbarn is the mean proton-proton inelastic cross section, $\eta_{\rm pp} = 0.5$ is the elasticity of pp collisions, Σ_g is the gas surface density of the galactic disc, c is the speed of light, D_0 is the diffusion coefficient at the galactic midplane, $\mu_{\rm p} = 1.17$ is the number density of nucleons per proton, and $m_{\rm H} = 1.67 \times 10^{-24}$ g is the mass of a hydrogen atom.

To evaluate the calorimetry fraction for a CR of energy E, we must therefore determine the midplane diffusion coefficient D_0 for CRs of that energy, which in turn depends on the streaming speed $V_{\rm st}$. This speed is dictated by the balance between excitation of the streaming instability and dissipation of the instability by ion-neutral damping, the dominant dissipation mechanism in the weakly-ionised neutral ISM. Balancing these two effects yields a CR proton streaming velocity

$$V_{\rm st} \approx \min\left[c, V_{\rm Ai}\left(1 + \frac{\gamma_d\,\chi\,M_A\,c\,\rho^{3/2}}{4\,\pi^{1/2}\,C\,e\,u_{\rm LA}\,\mu_i\,\gamma^{-p+1}}\right)\right] \tag{6}$$

where $V_{\rm Ai}$ is the ion Alfvén speed, $\gamma_d = 4.9 \times 10^{13}$ cm^3 g^{-1} is the ion-neutral drag coefficient, χ is the ionised mass fraction, $\rho = \Sigma_g / 2h_g$ is the midplane mass density of the ISM, C is the midplane number density of CRs, e is the elementary charge, $u_{\rm LA}$ is the velocity dispersion of Alfvén modes in the ISM at the outer scale of the turbulence, μ_i is the atomic mass of the dominant ion species, p is the index of the CR energy distribution, and $\gamma = E/m_p c^2$ is the Lorentz factor of the CR. Since γ-ray production in our model is dominated by galaxies with high star formation rates and gas surface densities within which i) the ISM is molecule-dominated, ii) the main ionised species is C^+, and iii) the magnetic field is set by a turbulent dynamo, we adopt the fiducial parameters of Ref. [11] appropriate for such galaxies. Specifically, we take $\chi = 10^{-4}$, $\mu_i = 12$, $M_A = 2$, $V_{\rm Ai} = u_{\rm LA}/\chi^{1/2} M_A$, and $u_{\rm LA} = \sigma_g/\sqrt{2}$, where σ_g is the gas velocity dispersion of the galaxy. We also adopt $p = 2.2$, consistent with our assumed injection spectrum.

At this point we have specified all the ingredients required to compute $f_{\rm cal}(E)$ for a galaxy of known Σ_g, σ_g, and h_g, save one: C, the CR number density. We estimate this as follows: consistent with our discussion in the main text, for a galaxy with star formation rate per unit area $\dot{\Sigma}_*$, the relativistic cosmic ray injection rate per unit area is

$$\frac{d\dot{N}}{dA} = \phi \dot{\Sigma}_* \int_{m_p c^2}^{\infty} (E/E_0)^{-p} e^{-E/E_{\rm cut}} dE \approx 5.87 \times 10^{42}\ \dot{\Sigma}_*\ {\rm s}^{-1} {\rm M}_\odot^{-1} {\rm yr}. \tag{7}$$

The CR number density at the midplane is then given by

$$C \approx \frac{t_{\rm loss}}{2h_g} \left(\frac{d\dot{N}}{dA} \right), \tag{8}$$

where $t_{\rm loss}$ is the CR loss time. This is given by $t_{\rm loss} = 1/\left(t_{\rm col}^{-1} + t_{\rm diff}^{-1}\right)$, where the timescale for losses in inelastic hadronic collisions is $t_{\rm col} = 1/\left(\rho \sigma_{\rm pp} \eta_{\rm pp} c/\mu_p m_{\rm H}\right)$ and the diffusive escape time $t_{\rm diff} = 100\ h_g^2 D_0^{-1}$, assuming CRs must escape from an extended region of height $\sim 10 h_g$. For the systems with high gas densities and high star formation rate that dominate γ-ray production, the loss time is generally dominated by collisional losses. Conversely, for systems forming stars more sedately and with lower density environments, it is generally determined by the diffusive escape time. (Here we have estimated the diffusion coefficient to first order by assuming CRs diffuse close to the ion Alfvén speed.)

Galaxy gas properties

Our calorimetry model requires, as input, the gas surface density Σ_g, scale height h_g, and velocity dispersion σ_g, along with the surface density of star formation $\dot{\Sigma}_*$. However, the CANDELS data set that we use provides only the cosmological redshift z, stellar mass M_*, half-light or effective radius $R_{\rm e}$ (corrected to 5000 Å according to Ref. [48]), and total star formation rate $\dot{M}_*$ for our sample galaxies. We must therefore estimate the gas properties from observed correlations between gas and stellar properties. We do so as follows.

The half light radius $R_{\rm e}$ at 5000 Å serves as a first order estimate of how the star formation and matter are distributed throughout the galactic disc. We therefore estimate the star formation rate surface density as $\dot{\Sigma}_* = \dot{M}_*/2\pi R_{\rm e}^2$ and the stellar surface density as $\Sigma_* = M_*/2\pi R_{\rm e}^2$. We estimate the gas surface density from the observed correlation between gas, stellar, and star formation surface densities given by Ref. [28]:

$$\frac{\Sigma_g}{{\rm M}_\odot\ {\rm pc}^{-2}} = 10^{10.28} \frac{\dot{\Sigma}_*}{{\rm M}_\odot\ {\rm yr}^{-1}\ {\rm pc}^{-2}} \left(\frac{\Sigma_*}{{\rm M}_\odot\ {\rm pc}^{-2}} \right)^{-0.48} \tag{9}$$

Similarly, there is a strong correlation between galaxy star formation rates and velocity dispersions, which we use to derive σ_g. For this purpose we fit the relationship using the MaNGA galaxy sample [29]. A powerlaw fit to the data obtained in this survey (Fig. 6 of [29]) gives

$$\frac{\sigma_g}{{\rm km\ s}^{-1}} = 32.063 \left(\frac{\dot{M}_*}{{\rm M}_\odot\ {\rm yr}^{-1}} \right)^{0.096} \tag{10}$$

Finally, we derive the gas scale height under the assumption that the gas is in vertical hydrostatic equilibrium, in which case the scale height is [30]

$$h_g = \frac{\sigma_g^2}{\pi G(\Sigma_g + \Sigma_*)} \tag{11}$$

When applying this model to the galaxies in the CANDELS sample, we take the stellar properties from Ref. [27] and correct the half-light radii to 5000 Å following the procedure outlined in Ref. [48]. This gives us a sufficient set of input parameters to apply our model.

References

1. Ackermann, M. *et al.* The Spectrum of Isotropic Diffuse Gamma-Ray Emission between 100 MeV and 820 GeV. **799,** 86. arXiv: 1410.3696 [astro-ph.HE] (Jan. 2015).
2. Fields, B. D., Pavlidou, V. & Prodanović, T. Cosmic Gamma-ray Background from Star-forming Galaxies. **722,** L199–L203. arXiv: 1003.3647 [astro-ph.CO] (Oct. 2010).
3. Ackermann, M. *et al.* GeV Observations of Star-forming Galaxies with the Fermi Large Area Telescope. **755,** 164. arXiv: 1206.1346 [astro-ph.HE] (Aug. 2012).
4. Xia, J.-Q., Cuoco, A., Branchini, E. & Viel, M. Tomography of the Fermi-LAT γ-Ray Diffuse Extragalactic Signal via Cross Correlations with Galaxy Catalogs. **217,** 15. arXiv: 1503.05918 [astro-ph.CO] (Mar. 2015).
5. Linden, T. Star-forming galaxies significantly contribute to the isotropic gamma-ray background. **96,** 083001. arXiv: 1612.03175 [astro-ph.HE] (Oct. 2017).
6. Ajello, M., Di Mauro, M., Paliya, V. S. & Garrappa, S. The γ-Ray Emission of Star-forming Galaxies. **894,** 88. arXiv: 2003.05493 [astro-ph.GA] (May 2020).
7. Cuoco, A., Komatsu, E. & Siegal-Gaskins, J. M. Joint anisotropy and source count constraints on the contribution of blazars to the diffuse gamma-ray background. **86,** 063004. arXiv: 1202.5309 [astro-ph.CO] (Sept. 2012).
8. Di Mauro, M., Calore, F., Donato, F., Ajello, M. & Latronico, L. Diffuse γ-Ray Emission from Misaligned Active Galactic Nuclei. **780,** 161. arXiv: 1304.0908 [astro-ph.HE] (Jan. 2014).
9. Qu, Y., Zeng, H. & Yan, D. Gamma-ray luminosity function of BL Lac objects and contribution to the extragalactic gamma-ray background. **490,** 758–765. arXiv: 1909.07542 [astro-ph.HE] (Nov. 2019).
10. Ajello, M. *et al.* The Origin of the Extragalactic Gamma-Ray Background and Implications for Dark Matter Annihilation. **800,** L27. arXiv: 1501.05301 [astro-ph.HE] (Feb. 2015).
11. Krumholz, M. R. *et al.* Cosmic ray transport in starburst galaxies. **493,** 2817–2833. arXiv: 1911.09774 [astro-ph.HE] (Apr. 2020).
12. Fornasa, M. & Sánchez-Conde, M. A. The nature of the Diffuse Gamma-Ray Background. **598,** 1–58. arXiv: 1502.02866 [astro-ph.CO] (Oct. 2015).
13. Bell, A. R. Cosmic ray acceleration. *Astroparticle Physics* **43,** 56–70 (Mar. 2013).
14. Woosley, S. E. & Weaver, T. A. *The evolution and explosion of massive Stars II: Explosive hydrodynamics and nucleosynthesis* NASA STI/Recon Technical Report N. Aug. 1995.
15. Dermer, C. D. & Powale, G. Gamma rays from cosmic rays in supernova remnants. **553,** A34. arXiv: 1210.8071 [astro-ph.HE] (May 2013).

16. Bell, A. R. The acceleration of cosmic rays in shock fronts - I. **182,** 147–156 (Jan. 1978).

17. Blandford, R. & Eichler, D. Particle acceleration at astrophysical shocks: A theory of cosmic ray origin. **154,** 1–75 (Oct. 1987).

18. Caprioli, D. Understanding hadronic gamma-ray emission from supernova remnants. **2011,** 026. arXiv: 1103.2624 [astro-ph.HE] (May 2011).

19. Caprioli, D. Cosmic-ray acceleration in supernova remnants: non-linear theory revised. **2012,** 038. arXiv: 1206.1360 [astro-ph.HE] (July 2012).

20. Chabrier, G. Galactic Stellar and Substellar Initial Mass Function. **115,** 763–795. arXiv: astro-ph/0304382 [astro-ph] (July 2003).

21. Heger, A., Fryer, C. L., Woosley, S. E., Langer, N. & Hartmann, D. H. How Massive Single Stars End Their Life. **591,** 288–300. arXiv: astro-ph/0212469 [astro-ph] (July 2003).

22. Kafexhiu, E., Aharonian, F., Taylor, A. M. & Vila, G. S. Parametrization of gamma-ray production cross sections for p p interactions in a broad proton energy range from the kinematic threshold to PeV energies. **90,** 123014. arXiv: 1406.7369 [astro-ph.HE] (Dec. 2014).

23. Razzaque, S., Mészáros, P. & Zhang, B. GeV and Higher Energy Photon Interactions in Gamma-Ray Burst Fireballs and Surroundings. **613,** 1072–1078. arXiv: astro-ph/0404076 [astro-ph] (Oct. 2004).

24. Franceschini, A. & Rodighiero, G. The extragalactic background light revisited and the cosmic photon-photon opacity. **603,** A34. arXiv: 1705.10256 [astro-ph.HE] (July 2017).

25. Berezinsky, V. & Kalashev, O. High-energy electromagnetic cascades in extragalactic space: Physics and features. **94,** 023007. arXiv: 1603.03989 [astro-ph.HE] (July 2016).

26. Grogin, N. A. *et al.* CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey. **197,** 35. arXiv: 1105.3753 [astro-ph.CO] (Dec. 2011).

27. van der Wel, A. *et al.* Structural Parameters of Galaxies in CANDELS. **203,** 24. arXiv: 1211.6954 [astro-ph.CO] (Dec. 2012).

28. Shi, Y. *et al.* Extended Schmidt Law: Role of Existing Stars in Current Star Formation. **733,** 87. arXiv: 1103.3711 [astro-ph.CO] (June 2011).

29. Yu, X. *et al.* What drives the velocity dispersion of ionized gas in star-forming galaxies? **486,** 4463–4472. arXiv: 1904.11431 [astro-ph.GA] (July 2019).

30. Forbes, J., Krumholz, M. & Burkert, A. Evolving Gravitationally Unstable Disks over Cosmic Time: Implications for Thick Disk Formation. **754,** 48. arXiv: 1112.1410 [astro-ph.GA] (July 2012).

31. Kennicutt Robert C., J. The Global Schmidt Law in Star-forming Galaxies. **498,** 541–552. arXiv: astro-ph/9712213 [astro-ph] (May 1998).

32. Crain, R. A., McCarthy, I. G., Frenk, C. S., Theuns, T. & Schaye, J. X-ray coronae in simulations of disc galaxy formation. **407,** 1403–1422. arXiv: 1005.1642 [astro-ph.CO] (Sept. 2010).

33. Kornecki, P. *et al.* γ-ray/infrared luminosity correlation of star-forming galaxies. **641,** A147. arXiv: 2007.07430 [astro-ph.HE] (Sept. 2020).

34. Bothwell, M. S. *et al.* The star formation rate distribution function of the local Universe. **415,** 1815–1826. arXiv: 1104.0929 [astro-ph.CO] (Aug. 2011).

35. Bothwell, M. S. *et al.* Erratum: The star formation rate distribution function of the local Universe. **438,** 3608–3608 (Mar. 2014).

36. Madau, P. & Dickinson, M. Cosmic Star-Formation History. **52,** 415–486. arXiv: 1403.0007 [astro-ph.CO] (Aug. 2014).

37. Ando, S., Fornasa, M., Fornengo, N., Regis, M. & Zechlin, H.-S. Astrophysical interpretation of the anisotropies in the unresolved gamma-ray background. **95,** 123006. arXiv: 1701.06988 [astro-ph.HE] (June 2017).

38. Manconi, S. *et al.* Testing gamma-ray models of blazars in the extragalactic sky. **101,** 103026. arXiv: 1912.01622 [astro-ph.HE] (May 2020).

39. Tamborra, I., Ando, S. & Murase, K. Star-forming galaxies as the origin of diffuse high-energy backgrounds: gamma-ray and neutrino connections, and implications for starburst history. **2014,** 043. arXiv: 1404.1189 [astro-ph.HE] (Sept. 2014).

40. Komis, I., Pavlidou, V. & Zezas, A. Extragalactic gamma-ray background from star-forming galaxies: Will empirical scalings suffice? **483,** 4020–4030. arXiv: 1711.11046 [astro-ph.HE] (Mar. 2019).

41. Ando, S. & Pavlidou, V. Imprint of galaxy clustering in the cosmic gamma-ray background. **400,** 2122–2127. arXiv: 0908.3890 [astro-ph.HE] (Dec. 2009).

42. Hooper, D., Linden, T. & Lopez, A. Radio galaxies dominate the high-energy diffuse gamma-ray background. **2016,** 019. arXiv: 1604.08505 [astro-ph.HE] (Aug. 2016).

43. Willott, C. J., Rawlings, S., Blundell, K. M., Lacy, M. & Eales, S. A. The radio luminosity function from the low-frequency 3CRR, 6CE and 7CRS complete samples. **322,** 536–552. arXiv: astro-ph/0010419 [astro-ph] (Apr. 2001).

44. Cherenkov Telescope Array Consortium *et al. Science with the Cherenkov Telescope Array* (2019).

45. di Sciascio, G. & Lhaaso Collaboration. The LHAASO experiment: From Gamma-Ray Astronomy to Cosmic Rays. *Nuclear and Particle Physics Proceedings* **279-281,** 166–173. arXiv: 1602.07600 [astro-ph.HE] (Oct. 2016).

46. Ballet, J., Burnett, T. H., Digel, S. W. & Lott, B. Fermi Large Area Telescope Fourth Source Catalog Data Release 2. *arXiv e-prints,* arXiv:2005.11208. arXiv: 2005.11208 [astro-ph.HE] (May 2020).

47. Peng, F.-K., Wang, X.-Y., Liu, R.-Y., Tang, Q.-W. & Wang, J.-F. First Detection of GeV Emission from an Ultraluminous Infrared Galaxy: Arp 220 as Seen with the Fermi Large Area Telescope. **821,** L20. arXiv: 1603.06355 [astro-ph.HE] (Apr. 2016).

48. van der Wel, A. *et al.* 3D-HST+CANDELS: The Evolution of the Galaxy Size-Mass Distribution since z = 3. **788,** 28. arXiv: 1404.2844 [astro-ph.GA] (June 2014).