Characterising the signatures of star-forming galaxies in the extra-galactic $\gamma$-ray background

Ellis R. Owen$^{1,2,3}$*, Khee-Gan Lee$^4$, Albert K. H. Kong$^1$

$^1$Institute of Astronomy, National Tsing Hua University, Hsinchu, Taiwan (ROC)
$^2$Center for Informatics and Computation in Astronomy, National Tsing Hua University, Hsinchu, Taiwan (ROC)
$^3$Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey, RH5 6NT, United Kingdom
$^4$Kavli Institute for the Physics and Mathematics of the Universe, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

ABSTRACT

Galaxies experiencing intense star-formation episodes are expected to be rich in energetic cosmic rays (CRs). These CRs undergo hadronic interactions with the interstellar gases of their host to drive $\gamma$-ray emission, which has already been detected from several nearby starbursts. Unresolved $\gamma$-ray emission from more distant star-forming galaxies (SFGs) is expected to contribute to the extra-galactic $\gamma$-ray background (EGB). However, despite the wealth of high-quality all-sky data from the Fermi-LAT $\gamma$-ray space telescope collected over more than a decade of operation, the exact contribution of such SFGs to the EGB remains unsettled. We investigate the high-energy $\gamma$-ray emission from SFGs up to redshift $z = 3$ above a GeV, and assess the contribution they can make to the EGB. We show the $\gamma$-ray emission spectrum from a SFG population can be determined from just a small number of key parameters, from which we model a range of possible EGB realisations. We demonstrate that populations of SFGs leave anisotropic signatures in the EGB, and that these can be accessed using the spatial power spectrum. Moreover, we show that such signatures will be accessible with ongoing operation of current $\gamma$-ray instruments, and detection prospects will be greatly improved by the next generation of $\gamma$-ray observatories, in particular the Cherenkov Telescope Array.

Keywords: cosmic rays – gamma-rays: diffuse background – gamma-rays: galaxies – galaxies: starburst – galaxies: star formation – galaxies: ISM

1 INTRODUCTION

The extra-galactic $\gamma$-ray background (EGB) has been measured by EGRET (Sreekumar et al. 1998; Strong et al. 2004) and Fermi-LAT (Ackermann et al. 2015) to be a power-law of spectral index of around $-2.3$ from 100 MeV up to an exponential cut-off at 300 GeV above which with detections persist to 820 GeV (Ackermann et al. 2015). The EGB can be decomposed into a component arising from the $\gamma$-ray emission of resolved extra-galactic sources, and a second component (sometimes referred to as the isotropic $\gamma$-ray background, IGRB) that emerges from the accumulation over redshift of all unresolved $\gamma$-ray emitting sources beyond our Galaxy, extending to the furthest reaches of the observable Universe. The physical origin of the unresolved component is thought to be a combination of unresolved active galactic nuclei (AGN) (e.g. Inoue 2011; Singal et al. 2012; Ajello et al. 2015) and star-forming galaxies (SFGs) (e.g. Bhattacharya & Sreekumar 2009; Fields et al. 2010; Lamastra et al. 2017). A further cascade component contributes as much as 50% of the flux below a TeV (Coppi & Aharonian 1997; Kneiske & Mannheim 2008; Kalashev et al. 2009; Berezhinskii et al. 2011; Wang et al. 2011; Inoue & Ioka 2012), which arises from high-energy $\gamma$-rays undergoing pair-production and subsequent Compton scattering in the extra-galactic background light, EBL (e.g. Madau & Phinney 1996).

The exact balance between the sources of the EGB remains unsettled. It has been argued that the majority of the flux originates in unresolved blazars (Inoue & Totani 2009; Singal et al. 2012; Ajello et al. 2015). These are complimented by radio galaxies (Inoue 2011; Di Mauro et al. 2014; Wang & Loeb 2016; Stecker et al. 2019), also being active galactic nuclei (AGN) but discriminated from blazars by their viewing angle (sometimes referred to as misaligned AGNs, or MAGNs), and flat spectrum radio quasars, FSRQs (Ajello et al. 2012) – see also (e.g. Stecker et al. 1993; Abdio et al. 2010b). Star-forming galaxies (SFGs) are also thought to make an important contribution, perhaps accounting for up to several tens of percent of the EGB intensity (Bhattacharya & Sreekumar 2009; Fields et al. 2010; Tamborra et al. 2014; Lamastra et al. 2017),1, but their exact origins have also been considered, including contributions from galaxy clusters (Zandanel et al. 2015), cascades from protons and/or heavy nuclei and their subsequent photo-disintegration/photo-pion production in cosmic radiation fields (Kalashev et al. 2009; Ahlers & Salvado

* E-mail: erowen@gapp.nthu.edu.tw (ERO)

© 2020 The Authors
contribution remains poorly constrained (e.g. Komis et al. 2019). Intense starburst episodes experienced by SFGs yield abundant stellar end-products soon after the onset of star-formation. The shocked violent astrophysical environments contributed by such end-products – in particular supernova explosions and their remnants – provide ample low-energy charged seed particles and magnetised shocks, which can boost the seeds to relativistic energies (e.g. by Fermi acceleration – Fermi 1949) to form a rich interstellar reservoir of CRs. Their containment within the internal environment of the host galaxy by a rapidly-amplified magnetic field (Schober et al. 2013; Owen et al. 2018) particularly enhances the CR density in SFGs, which may undergo pion-producing hadronic interactions with the ambient gases of their host to drive γ-ray emission (e.g. Pfrommer et al. 2017). Such emission has already been detected from several nearby starbursts, with several nearby SFGs having already been resolved in γ-ray observations with Fermi–LAT – M82 and NGC 253 (Acero et al. 2009; VERITAS Collaboration 2009; Abdo et al. 2010a; Rephaeli & Persic 2014), NGC 2146 (Tang et al. 2014), Arp 220 (Peng et al. 2016; Griffin et al. 2016; Yoast-Hull et al. 2017), M33 and Arp 299 (Xi et al. 2020; Ajello et al. 2020), NGC 2403 and NGC 3424 (Peng et al. 2019; Ajello et al. 2020), and NGC 1068 and NGC 4945 (Ackermann et al. 2012) – of these, some (M82 and NGC 253) also have higher energy counterpart detections with VERITAS and H.E.S.S. (Acero et al. 2009; Karlsson & for the VERITAS collaboration 2009; H. E. S. S. Collaboration et al. 2018).

The superposition of sources in the EGB must include contributions over a broad range of redshifts, particularly at lower energies (below around 30 GeV) where the cascade effect is less severe (e.g. Gilmore et al. 2009). The cosmological propagation of γ-rays from their source must therefore be considered when modelling the formation of the EGB. At high energies, γ-ray photons interact with soft EBL photons and the cosmological microwave background (CMB) radiation to produce a cascade of secondary charged leptons via pair-production (Heitler 1954; Madau & Phinney 1996). These secondary leptons can cool via Compton scattering (and, to a lesser degree, synchrotron emission in intergalactic magnetic fields) to form the diffuse secondary flux of γ-rays (e.g. Wang et al. 2011). Practically, this leads to an attenuation effect arising for higher energy γ-rays as they propagate, but an emission effect at lower energies along our line of sight to a source. The resulting energy spectrum can become heavily distorted over large distances – particularly from sources located at higher redshifts, when intergalactic radiation fields would have been more intense than today and cascade losses more severe, and this must be carefully modelled. This process is of often valuable to researchers, as the attenuation of high-energy γ-rays from distant sources can be used as a tool to probe and constrain EBL radiation fields which, in turn, can offer clues to the formation and evolution of galaxy populations over redshift (Mazin & Raue 2007; Fermi–LAT Collaboration 2018). But in this work, we are concerned with the γ-rays themselves, and the modification and confluence of the energy spectra of a broad distribution of sources located at different redshifts. The attenuation of γ-rays is more severe at higher energies, and does not operate below ~ 0.1 GeV (energies for which the γ-ray path length would presumably extend to distances comparable to the size of the visible Universe – see, e.g. Neronov & Semikoz 2012).

The Universe remains relatively transparent to z > 5 for γ-ray energies of up to around 30 GeV (Gilmore et al. 2009) but becomes optically thick by z < 0.1 in the TeV band (Franceschini et al. 2008; Dominguez et al. 2011; Gilmore et al. 2012; Inoue et al. 2013a). This presents a challenge for spectral studies, where higher energy emission would become inaccessible over relatively short cosmological distances. However, the CR spectrum of SFGs peaks at around 1–10 GeV and, given the weak energy-dependence of the pp-interaction cross section above its interaction threshold (Kafexhiu et al. 2014), their γ-ray emission would presumably reflect this and much of their flux would fall below the energies most strongly affected by EBL attenuation. The peak of cosmic star-formation (and presumably the redshift at which SFGs are most abundant) arises at approximately z ∼ 2 (Madau & Dickinson 2014; Fermi–LAT Collaboration 2018) – a distance from which EBL transmittance at 10s of GeV is reasonable. We therefore expect that SFGs would make a substantial contribution to the observed EGB at these energies.

Studies of the γ-ray background with Fermi–LAT have recently revealed small-scale anisotropies between 0.5 and 500 GeV (Fornasa et al. 2016; Ackermann et al. 2018), thought to arise from two distinct source populations (SFGs and BL Lac. objects). In this work, we investigate such signatures imprinted by a SFG contribution to the γ-ray background and assess their form and its sensitivity to the underlying source population’s basic characteristics. We consider a model for the emission, attenuation, reprocessing and cosmological propagation of γ-rays, also accounting for the redshift-evolution of their star-forming galaxy populations. We consider the characteristic separation of galaxies at a given redshift, described by the galaxy power spectrum (Tegmark et al. 2004), and compute how the resulting redshift-integrated length-scale would become imprinted into the γ-ray background. We demonstrate how the resulting angular power spectrum of the anisotropies in the EGB is sensitive to the properties and redshift evolution of the source galaxy population, and consider the observational prospects of such a signature with ongoing Fermi–LAT observations, and the up-coming Cherenkov Telescope Array (CTA).

This paper is organised as follows. In Section 2, we outline the CR interactions and discuss the γ-ray emission from SFGs. We consider the relevant pair-production processes arising within a galactic environment, and assess the impacts these have on the emitted γ-ray spectrum from a SFG. In Section 3, we consider the properties of SFG populations, model the cosmological propagation of γ-rays and introduce our EGB anisotropy model. We present our results in Section 4 and show the relation between EGB anisotropy signatures and key parameters in our model. We also consider the detection prospects of such signatures. We provide a brief summary and draw conclusions in Section 5.

2 γ-RAY EMISSION FROM STAR-FORMING GALAXIES

We adopt the following notation convention. We express particle energies in terms of their Lorentz factor, e.g. for protons, the total energy $E_p = \gamma_p m_p c^2$, which is related to the proton kinetic energy by $T_p = (\gamma_p - 1)m_p c^2$. Photon energies (including γ-rays) are expressed in terms of the dimensionless quantity $\epsilon_\gamma = E_\gamma/m_e c^2$, i.e. normalised to the electron rest mass energy.
2.1 Cosmic ray interactions

Energetic hadronic cosmic rays may interact through various channels, however the internal conditions of typical star-forming galaxies would strongly favour proton-proton (hereafter pp) pion-production mechanisms (Owen et al. 2018, 2019b). Hadronic interactions with radiation fields (photo-pion or photo-pair production) are comparatively unimportant, despite of the intense radiation fields generated by the young stellar populations (Owen et al. 2018). The pp interaction can arise above a threshold proton kinetic energy of $T_p^{th}/c^2 = 2m_p c^2 + m_n^2 c^4/2m_p = 0.28 \text{ GeV}/c^2$, and involves a two-stage mechanism: the first step is the formation of resonance baryons via $pp \rightarrow p\Delta^+$ or $pp \rightarrow n\Delta^+$ (Almeida et al. 1968; Skorodko et al. 2008), while the second step is their decay (on timescales of $5.63 \times 10^{-24} \text{ s}$ – see Patrignani et al. 2016) according to

$$\Delta^+ \rightarrow \begin{cases} \text{p}\pi^0\xi_0(\pi^0)\xi_\pm(\pi^+\pi^-) , \text{ (1)} \\
\text{p}\pi^0\xi_0(\pi^0)\xi_\pm(\pi^+\pi^-) , \text{ or (2)} \\
n\pi^0\xi_0(\pi^0)\xi_\pm(\pi^+\pi^-) , \text{ for which the terms } \xi_0 \text{ and } \xi_\pm \text{ are the energy-dependent multiplicities of the three pion species (Jain & Santra 1993; Lebedevic 2014). Despite the energy-dependence of the multiplicities, the overall intrinsic production rate of each pion species is relatively energy-independent, with ratios of $\{\pi^+, \pi^-, \pi^0\} = \{0.6, 0.1, 0.3\}$ arising at 1 GeV, and slowly varying to $\{0.3, 0.4, 0.3\}$ by 50 GeV, with negligible evolution thereafter (Jacobsen 2015). $\gamma$-ray production proceeds (with a branching ratio of 99.8%) through neutral pion decays $\pi^0 \rightarrow 2\gamma$ on a timescale of $8.5 \times 10^{-17} \text{ s}$ (Tanabashi et al. 2018) and, given the weak energy-dependence of the pp interaction cross-section and the $\pi^0$ formation multiplicity, would yield a $\gamma$-ray spectrum closely tracing the shape of the proton spectrum driving the emission. $\gamma$-ray emission can also arise from inverse Compton scattering of secondary electrons injected by charged pion decays. However, the resulting emissivity by this channel is not expected to dominate at energies above 0.1-1 GeV (e.g. Chakraborty & Fields 2013; Pfrommer et al. 2017) and so is not included.

2.1.1 Hadronic interaction rate

The volumetric rate at which pp interactions arise is given by

$$n_{pp}(\gamma_p) = n_{HI}(\gamma_p) n_p(\gamma_p) c \sigma_{pp}(\gamma_p) ,$$

(3)

where $n_{HI}(\gamma_p)$ is the average ambient gas density within the host galaxy, $n_p$ is the CR proton density, and $\sigma_{pp}$ is the total inelastic pp interaction cross section, which may be parameterised as

$$\sigma_{pp} = \left[30.7 - 0.96 \ln(\chi) + 0.18(\ln(\chi))^2\right] \left[1 - \chi^{-1.9}\right]^{-3} \text{ mb} ,$$

(4)

(Kafexhiu et al. 2014), where $\chi = T_p/T_p^{th}$, for $T_p^{th} = (\gamma_p^{th} - 1)m_{p}\text{c}^2$ is the threshold proton kinetic energy, below which the interaction cannot occur. Equation 4 therefore represents the volumetric loss rate of CR protons due to the pp process within the interstellar medium (ISM) of the host starburst galaxy. This is different (although related) to the production rate of $\gamma$-rays, which relies on the formation and subsequent decay of neutral pions. The differential $\gamma$-ray inclusive cross section by the $pp \rightarrow pp\pi^0$ interaction channel may be written as

$$\frac{d\sigma_{pp}(\gamma_p, \epsilon_\gamma)}{d\epsilon_\gamma} = \mathcal{P}(\gamma_p) \times \mathcal{F}(\gamma_p, \epsilon_\gamma) ,$$

(5)

where the peak function $\mathcal{P}$ and spectral form $\mathcal{F}$ are also well-parametrised to an accuracy of better than 10 per cent by Kafexhiu et al. (2014).

2.1.2 $\gamma$-ray production in starburst galaxies

We compute the $\gamma$-ray spectral emissivity from the CR proton density, $n_p(\gamma_p)$ (see section 2.2; note that this is also a differential quantity such that $n_p(\gamma_p) d\gamma_p$ is the number density of CR protons in the energy interval $d\gamma_p$), and the inclusive differential $\gamma$-ray production cross-section.2 The spectral emissivity of $\gamma$-rays may be written as

$$\frac{d\mathcal{L}(\epsilon_\gamma)}{d\epsilon_\gamma} = c \int_{\gamma_p^{th}}^{\gamma_p} \frac{\mathcal{P}(\gamma_p) \mathcal{F}(\gamma_p, \epsilon_\gamma)}{d\epsilon_\gamma} n_p(\gamma_p) d\gamma_p ,$$

(6)

where $c$ is the speed of light, and where we set $\gamma_p^{th} = 50 \text{ PeV}/m_{p}\text{c}^2$ as the upper limit for the acceleration of CR protons in starbursts (as argued in Peretti et al. 2019).

2.2 Cosmic ray spectrum and energy budget

To compute the $\gamma$-ray spectral emissivity of a galaxy (equation 6), we require knowledge of the internal CR spectrum. Typically, this is well-described by a simple power-law, with a spectral index $\Gamma$ between -2.1 and -2.7 depending on the exact environment (Kotera & Olinto 2011). In star-forming regions, the particle spectrum is presumably freshly accelerated and would be described by a less-steep spectral index, with more CRs at higher energies. Here, we initially adopt a proton spectrum of $\Gamma = -2.1$, being similar to the characteristic value (of between -1.9 and -2.3) inferred for local SFGs detected in $\gamma$-rays (see, e.g. Tamborra et al. 2014; Rojas-Bravo & Araña 2016; Ajello et al. 2020).3 This choice of index is also comparable to regions within the Milky Way CRs are thought to be freshly accelerated, i.e. towards the galactic ridge (see, e.g. Allard et al. 2007; Kotera et al. 2010). We relax this choice later, in section 4, where we consider alternative CR index values within this range.

2.2.1 Cosmic ray luminosity

We estimate the CR proton density within a SFG using equation A8 (see Appendix A for details), where we consider that the CR proton spectrum and density can be parameterised by just four quantities: the star-formation rate of the host galaxy, $\Phi_{SF}$, the size of the nuclear starburst region, $R$, the CR spectral index, $\Gamma$, and the maximum CR proton energy, $E_{max}$. We argue that there is insufficient motivation to consider substantial variation of other quantities, such as those pertaining to the CR diffusion coefficient (equation A3), ISM

---

2 Other works (e.g. Peretti et al. 2020) instead consider the $\gamma$-ray flux density of M82 as a prototype, and model the $\gamma$-ray emission of other SFGs by scaling this with star-formation rate.

3 While spectral indices are observed for the $\gamma$-ray emission from nearby SFGs, it is expected that the $\gamma$-ray spectrum above ~1 GeV would closely follow the CR proton spectrum, because of the limited energy-dependence of the inclusive $\pi^0$ formation cross section – e.g. Kafexhiu et al. (2014).
density/structure, or the characteristic fraction of CRs advected by galactic outflows, and we fix these are their values stated in Appendix A.

2.2.2 Internal attenuation of $\gamma$-rays

While the production of high-energy $\gamma$-rays within SFGs is predominantly regulated by the hadronic pp interactions of CRs (see Figure 1), their resulting $\gamma$-ray emission spectra is more complicated than this would imply. $\gamma$-ray absorption processes would operate within the internal environment of a SFG, substantially modifying the ensuing emitted spectrum. Recent works (e.g. Vereen & de Vries 2020) have considered that sufficiently dense gas clouds within the ISM of a host galaxy could significantly attenuate $\gamma$-rays through pair-production on the dense gas. This is compelling, as such dense clouds would act as a CR beam dump via the pp-interaction, making these the principal sites of $\gamma$-ray production by this channel. However, such a mechanism across a galaxy would require very substantial gas column densities along many ISM lines of sight. This would imply a heavy loading of the ISM with large, dense clouds. While this cannot be ruled out – and may be one of several processes operating to modify the $\gamma$-ray spectrum of a galaxy on a global scale – in this work we instead consider the $\gamma$-ray attenuation that would result from ambient radiation fields (in particular, those due to the stellar population and the component of the stellar radiation reprocessed starlight, proceeds as:

$$\gamma + \gamma \rightarrow e^+ + e^-,$$

(7)

at a rate of

$$N_{\gamma\gamma}(\epsilon) = c \int_0^\infty d\epsilon_{\text{ph}}(\epsilon) \sigma_{\gamma\gamma}(\epsilon),$$

(8)

where $n_{\text{ph}}$ is the spectral number density of target photons, $\sigma_{\gamma\gamma}$ is the $\gamma\gamma$ interaction cross section (see, e.g. Gould 2005) and $\epsilon_c \approx \epsilon_{\gamma}/2$ is the invariant interaction energy for an isotropic radiation field. The electron pairs formed in this process can predominantly cool by Compton up-scattering photons in ambient radiation fields, or thermalise in the interstellar gas depending on their energy (with other processes arising at a lower rate – see Owen et al. 2018 for a comparison of various cooling timescales experienced by electrons in typical SFGs). Electrons below $\sim$ 100 MeV predominantly thermalise in less than 1 Myr in ISM conditions (Owen et al. 2018), corresponding to a diffusive length-scale of $\sim$ 1 kpc. However, most electrons are injected at much higher energies than this, above 10s of GeV (reflecting the energies where $\gamma$-ray attenuation is strongest – see Figure 1). For these, thermalisation timescales are much longer so electromagnetic cascades tend to develop instead, where electrons Compton up-scatter ambient interstellar radiation field (ISRF) photons to high-energies (typically to form so-called secondary $\gamma$-rays or X-rays; see, e.g. Chakraborty & Fields 2013). At 100 GeV, for instance, electrons would thermalise over $\sim$ 1 Gyr in the typical SFG environment considered in Owen et al. 2018, while their Compton scattering timescale would be just a few kyr. The up-scattered secondary photons may undergo further pair-production, if they are of sufficient energy. Alternatively they may escape from the host galaxy, modifying the emitted $\gamma$-ray spectrum from the SFG. However, given that the majority of the secondary $\gamma$-ray emission develops from the attenuation of the highest energy primary $\gamma$-rays (cf. Figure 1), for which the flux is lowest (due to the power-law nature of the emissivity), their contribution to the emitted spectrum would be negligible (Fitoussi et al. 2017). As such, we argue that the final emitted $\gamma$-ray spectrum of a SFG can be well-described by the $\gamma$-ray emissivity from hadronic interactions, modified by their attenuation through pair production. The negligible secondary cascade emission from within the ISM of the host galaxy is not included in our model.

Without loss of generality, we define the characteristic $\gamma$-ray path length in a radiation field as $\ell_{\gamma\gamma}(\epsilon_\gamma, x) = c/N_{\gamma\gamma}(\epsilon_\gamma, x)$. This is the distance over which an interaction would typically arise under conditions specified at location $x$. It can be associated with a pair-production $\gamma$-ray optical thickness by

$$\tau_{\gamma\gamma}(\epsilon_\gamma, x) = \int_0^x d\epsilon' \ell_{\gamma\gamma}^{-1}(\epsilon_\gamma, \epsilon'),$$

(9)

over some path length $x$. In an isotropic black-body radiation field, this may be expressed as

$$\tau_{\gamma\gamma}^{BB}(\epsilon_\gamma, x) = \frac{2\pi\epsilon_\gamma^2}{\lambda_C} \int_0^x \Theta(x') \mathcal{J}(\epsilon_\gamma, x') d\epsilon' .$$

(10)

(Gould & Schréder 1967; Brown et al. 1973; Dermer & Menon 2009). Here, $\alpha_\ell$ is introduced as the fine structure constant, $\lambda_C$ is the electron Compton wavelength, $\Theta(x) = k_B T(x)/m_e c^2$, with $m_e$ as the electron rest mass, $c$ as the speed of light and $k_B$ is the Boltzmann constant, defines the dimensionless temperature at some position $x$, and the function $\mathcal{J}(\epsilon, x)$ is given by:

$$\mathcal{J}(\epsilon_\gamma, x) = \frac{1}{2\pi\epsilon_\gamma^2} \int_{1/\epsilon_\gamma\Theta(x)}^\infty \frac{\psi(\epsilon_\gamma, \Theta(x))}{\exp(y) - 1} dy$$

(11)

(e.g. Zdziarski & Svensson 1989), where the term

$$\psi(\epsilon_\gamma) = \frac{2}{\pi\epsilon_\gamma^2} \int_0^\epsilon \left( \frac{d\epsilon_\gamma}{\sigma B(\epsilon_\gamma)} \right)$$

(12)

specifies the change in the scattering cross-section compared to the (classical) Thomson cross-section $\sigma_T = \pi\epsilon_\gamma^2$, $\epsilon_\gamma = \epsilon_{\gamma\gamma}$ is the combined photon energy, and $r_\gamma$ is the classical electron radius. This is evaluated in Gould & Schréder 1967 (see also Brown et al. 1973). Equation 10 can be used to quantify the $\gamma$-ray attenuation factor within the host galaxy. Along a single line of sight $s$, this would simply be

$$\mathcal{A}(\epsilon, s) = \exp\left[-\tau_{\gamma\gamma}(\ell_{\gamma\gamma}, s, \epsilon_\gamma)\right] ,$$

(13)

however, when averaged through an extended SFG source (modelled as a uniformly attenuating sphere of radius $R$), we instead adopt an approximate characteristic attenuation factor specified by the size of the nucleus $R$ and the effective path length of the $\gamma$-rays, $\ell_{\gamma\gamma}$:

$$\mathcal{A}(\xi) = \exp\left[-\xi^{-2}\right] ,$$

(14)

If each of the secondary electrons adopts half of the energy of the primary $\gamma$-ray, the resulting Compton-scattered secondary $\gamma$-rays would have a peak energy of $E_{\gamma\gamma,i} = \left(E_{\gamma\gamma, i}/m_e c^2\right)^{1/2} E_{\text{peak}}$, where $E_{\text{peak}}$ is the peak energy of the target radiation field, and $E_{\gamma\gamma, i}$ for $i = \{1, 2\}$ are the primary and peak secondary $\gamma$-ray energies respectively. For a 100 GeV primary $\gamma$-ray, the characteristic secondary energy would be of order $\sim$ 10 MeV, while for a 100 TeV primary, it would be of order $\sim$ 10 TeV.
(see Appendix C for details) where \( \zeta = (R/\ell_{\gamma y})^{1/2} = r^{1/2} (\epsilon_y, R) \). We consider that \( \gamma \)-ray attenuation within a SFG is dominated by three radiation fields: (1) the CMB; (2) the ISRF from stars, and (3) the re-processed ISRF by interstellar dust. These may each be described by a Planck function of the form

\[
\eta_{ph} (\epsilon; \Theta) = \frac{8\pi \epsilon}{k} \exp \left( \frac{\epsilon}{\Theta} \right) - 1 ,
\]

where \( \Theta = \frac{k T_{ph}}{m_e c^2} \) is the dimensionless temperature of the radiation field, and \( f_{ph} \) is the dilution factor for geometrically distributed sources. The CMB is an undiluted radiation field, so \( f_{ph} \) is 1. Its temperature is a function of redshift, and so is given by \( T_{CMB}(z) = T_0 (1 + z) \), where \( T_0 = 2.73 \) K (Planck Collaboration et al. 2020) is the temperature of the CMB today. As a baseline model in the following results and calculations, we adopt a redshift of \( z = 2 \) (unless specified otherwise). The ISRF components due to stars and re-processed emission by dust would be a diluted blackbody, as the emission originates from the stars. The dilution factor in these two cases can be determined from the photon density in each radiation field. In general, this may be estimated as

\[
f_{ph} = \frac{L_C^3}{96 \pi^2 R^2 m_e c^3 \Theta (3) \zeta (3)} ,
\]

which is the ratio of the estimated photon density from the diluted black-body radiation field, compared to that expected for an undiluted black-body of the same temperature. Here, \( L \) is the total luminosity of the sources, \( \gamma (\ldots) \) is the gamma function, and \( \zeta (\ldots) \) is the Riemann zeta function. As a baseline choice, we adopt a characteristic value of \( R = 0.1 \) kpc for the size of a SFG nucleus, being comparable to the size of nearby starburst, for example NGC 253 (Weaver et al. 2002), for which \( R \approx 0.1 \) kpc, or M82 (found to have a core of 0.25 kpc – see de Grijse et al. 2001).

Young stars dominate the radiative emission from the stellar population of a SFG, and the dust optical depths are so great that a large fraction of the bolometric SFG luminosity is re-processed and re-radiated to IR wavelengths (Kennicutt 1998b). We consider that the total dust-reprocessed luminosity is comparable to the luminosity integrated over the full mid and far IR spectrum (8-1000\( \mu \)m). For SFGs, most of this will fall within the 10-120 \( \mu \)m spectral band (Kennicutt 1998a). As such, the luminosity of the dust emission from a SFG, \( L = L_{dust} \), is strongly coupled to its \( R_{SF} \), via:

\[
L_{IR} = 2.2 \times 10^{43} \left( \frac{R_{SF}}{1 M_\odot} \right) \text{erg s}^{-1}
\]

(Kennicutt 1998b), which is derived by applying the models of (Leitherer & Heckman 1995) for continuous starburst episodes of age 10-100 Myr, and a Salpeter (1955) initial stellar mass function between 0.1 and 30 \( M_\odot \). Presumably, this scaling relation would not be strongly sensitive to the exact choice of lower or upper IMF mass cut-off, if less than \( \sim 1 M_\odot \) or above 30 \( M_\odot \), for which the luminosity or number of stars (respectively) would not be substantial. We adopt this relation, which holds for the vast majority of SFGs (Bergvall et al. 2016), where the star-forming burst durations do not greatly exceed 100 Myr (Kennicutt 1998a). Interstellar dust emission is typically dominated by that from large grains (> 0.01\( \mu \)m), which are in thermal equilibrium with ambient interstellar radiation (e.g. Desert et al. 1990). The temperature of their emission \( T_{dust} \) (encoded by \( \Theta = \Theta_{dust} = k_B T_{dust}/m_e c^2 \)) in SFGs typically takes a characteristic value of a few tens of K. There is evidence for some redshift evolution (Magdis et al. 2012; Magnelli et al. 2014; Béthermin et al. 2015), with effective temperatures increasing from typical values of around 25 K at \( z = 0 \), to around 40 K by \( z = 4 \) (e.g. Schreiber et al. 2017, 2018). We adopt the empirical power law of Schreiber et al. (2018) to model this,

\[
T_{dust} = [32.9 \pm 2.4 + (z - 2)(4.60 \pm 0.35)] \text{K} ,
\]

where uncertainties are propagated through our model. Our fiducial model considers a redshift of \( z = 2 \), near the peak of cosmic star-formation (Madau & Dickinson 2014). This gives a corresponding dust temperature of 32.9 \( \pm \) 2.4 K. We model the dust-reprocessed radiation field to be spatially homogeneous and isotropic within the host SFG nucleus (up to a radius of \( R \)). The impact of detailed interstellar variations of the dust emission within SFGs (e.g. the clumpy distributions found in Bassett et al. 2017) is left to future work.

The total stellar radiative output power of stars in an SFG, \( L = L^* \), is dominated by young, massive, O and B type stars. It can be estimated from the dust luminosity \( L_{dust} \), using:

\[
L^* = (1 - \eta) L_{dust} \frac{f_{dust}}{0.4 - 0.2 f_{dust} + 0.6 \beta} ,
\]

(Inoue et al. 2000), where \( f_{dust} = 0.26 \) is the fraction of ionising stellar photons absorbed by ISM Hydrogen (from Petrosian et al. 1972, which derives the value from the Orion nebula), \( \beta = 0.6 \) is the averaged dust-absorption efficiency of non-ionising photons from central sources in ionised, star-forming regions (from Savage & Mathis 1979, which uses the extinction curve of the Galaxy), and \( \eta = 0.5 \) is the fraction of IR emission attributed to diffuse ISM gas, being distinct from the emission from star-forming regions (Helo 1986). This approach is valid for both strong starbursts (which emit almost all of their energy in IR – see Soifer et al. 1987) as well as moderate starbursts (where a large fraction of the stellar radiation may not be reprocessed by dust – see Buat & Xu 1996). We set the temperature of this radiation field to be \( T^* = 3 \times 10^4 \) K, to reflect the temperature of the dominant source population of massive O and B type stars.

The attenuative effects of these radiation fields on the emitted spectrum are demonstrated in Figure 1 for our fiducial case (\( R_{SF} = 10 M_\odot \text{yr}^{-1}, R = 0.1 \) kpc and at \( z = 2 \)). This demonstrates the severe impact of interstellar dust, which completely attenuates \( \gamma \)-rays above 10 TeV in this case. The CMB has some impact on lower-energy \( \gamma \)-rays, and would become more severely attenuating at higher redshifts. The (un-processed) starlight is comparatively unimportant, with a large fraction of the stellar emission having been reprocessed to IR wavelengths by interstellar dust.

### 2.3 Cosmological propagation and reprocessing of \( \gamma \)-rays

To form the \( \gamma \)-ray background as we observe it from Earth at \( z = 0 \), high energy photons emitted from source populations must propagate through intergalactic space over cosmological distances. During their propagation, \( \gamma \)-ray photons interact with soft EBL photons at IR and optical wavelengths. Fundamentally, this is the same process as that which leads to the attenuative losses of \( \gamma \)-rays within a SFG (considered in section 2.2.2). However, after their initial formation via pair-production (Heitler 1954; Madau & Phinney 1996), they Compton up-scatter EBL and CMB photons to form the diffuse secondary flux of \( \gamma \)-rays (e.g. Wang et al. 2011; Inoue et al. 2013a; Lacki et al. 2014) which continues to propagate and interact if photon energies remain sufficient. This cascade reprocessing is coupled with the concurrent red-shifting of the \( \gamma \)-ray beam, which
can be modelled using a covariant radiative transfer approach. We differentiate between dimensionless energies for soft EBL or CMB photons and γ-ray photons using the notation $\epsilon$ and $\epsilon_\gamma$ respectively. Moreover, $\gamma_e = E_e/m_e c^2$ is introduced as the electron Lorentz factor.

2.3.1 Cosmological γ-ray radiative transfer

The propagation of γ-rays through soft intergalactic radiation fields may be modelled using a radiative transfer approach, where γ-ray emitting populations form the source function, while the cascade process effectively operates as an attenuation process at high energies, and an emission process at lower energies. Over cosmological distances, the radiative transfer equation, in terms of redshift, takes the form:

$$\frac{dI_\gamma}{dz} = (1 + z) \left[ -\alpha_{\gamma\gamma} I_\gamma + \frac{j_\gamma}{\nu^3} \right] \frac{ds}{dz}$$

(20)

(e.g. Chan et al. 2019) where all quantities are Lorentz invariant, i.e. $I_\gamma = I_\nu/\nu^3$ for $I_\nu$ as the local ‘proper’ intensity (such that, in practice, co-moving absorption $\alpha_{\gamma\gamma}$ and emission $j_\gamma$ functions are used for the attenuation and cascade re-emission of γ-rays respectively, as well as co-moving frequency $\nu$), and $ds/dz$ for a flat Friedmann-Robertson-Walker (FRW) Universe is given by

$$\frac{ds}{dz} = \frac{c}{H_0(1+z)} \left( \Omega_{\gamma,0}(1+z)^4 + \Omega_{\Lambda,0}(1+z)^3 + \Omega_k,0 \right)^{-1/2}$$

(21)

(see, e.g. Peacock 1999), where $\Omega_{\Lambda,0} = 0.315 \pm 0.007$, $\Omega_k,0 = 0$ and $\Omega_{\gamma,0} = 0.685 \pm 0.007$ are the normalised density parameters for matter, radiation and dark energy respectively, and $H_0 = 100 h$ km s$^{-1}$ Mpc$^{-1}$ is the present value of the Hubble constant, where $h = 0.673 \pm 0.006$ (Planck Collaboration et al. 2020). We solve equation 20 by discretising a source SFG population into redshift shells. The discretised solutions are then integrated over redshift, to find the total EGB intensity from $z = 0$ to a maximum redshift, $z = z_{\text{max}} = 3$, a range which covers the peak of cosmic star-formation (Madau & Dickinson 2014) and would presumably account for the majority of γ-ray emission from SFGs.

2.3.2 γ-ray absorption and cascade reprocessing

The absorption of γ-rays by cascade pair-production in the EBL can be characterised by an absorption coefficient,

$$\alpha_{\gamma\gamma}(\epsilon, \epsilon_\gamma) = \frac{\sigma_T}{\epsilon} \int_0^\infty \epsilon_\gamma d\epsilon_\gamma' \frac{1}{\epsilon_\gamma'} \int_0^{\epsilon_\gamma'} d\epsilon \frac{d\epsilon}{d\epsilon_\gamma'} \psi(\epsilon_\gamma')$$

(22)

(Nikishov 1961; see also Gould & Schréder 1967; Brown et al. 1973), where $\psi(\epsilon_\gamma^*)$ retains its earlier definition of $e\epsilon_\gamma$, and $\epsilon_\gamma^*$ is given by equation 12. Over cosmological distances, a corresponding γ-ray optical depth due to pair-production may be written as

$$\tau_{\gamma\gamma}(z_\gamma, \epsilon_\gamma) = \int_0^{\epsilon_\gamma} \frac{\alpha_{\gamma\gamma}(\epsilon, \epsilon_\gamma)}{\epsilon_\gamma} \frac{ds}{dz} dz'$$

(23)

The EBL and its impact on γ-ray absorption has been extensively studied (see Dwek & Krennrich 2013; Cooray 2016 for reviews) via direct measurements in UV/optical and/or near-IR bands (e.g. Matsuka et al. 2011; Berta et al. 2011; Béthermin et al. 2012; Driver et al. 2016; Andrews et al. 2018), indirect measurements using the attenuation of high-energy γ-rays from extra-galactic sources (e.g. Desai et al. 2019; Abeyesekara et al. 2019; Paeschel 2019; Acciari et al. 2019), and theoretical models.

EBL models typically follow one of three approaches: (1) forward-evolutionary models, which convolve spectral models with cosmic star-formation histories to estimate the EBL’s development over redshift (e.g. Kneiske & Dole 2010; Finke et al. 2010); (2) backward-evolutionary models, which extrapolate observed properties of galaxies in the local Universe to higher redshifts (e.g. Franceschini et al. 2008; Domínguez et al. 2011; Helgason & Kashlinsky 2012; Stecker et al. 2012; Franceschini & Rodighiero 2017), and (3) semi-analytical models, SAMs (see Kauffmann & White 1993; Cole et al. 1994) of hierarchical galaxy formation (e.g. Gilmore et al. 2009; Younger & Hopkins 2011; Gilmore et al. 2012; Inoue et al. 2013a). Of these, forward-evolutionary models (category 1) suffer from several drawbacks. Notably, they do not trace the detailed evolution of crucial quantities which can impact the EBL spectrum, they are not able to reproduce certain observables (e.g. the observed
The star-forming galaxy contribution to the EGB

3 POPULATIONS OF STAR-FORMING GALAXIES

In section 2.2, it was shown that the γ-ray luminosity of a galaxy can be largely specified by its supernova (SN) event rate, $\psi_{\text{SN}}$. It was shown that this is directly related to the star-formation rate, $\psi_{\text{SF}}$ if assuming an IMF ($\psi_{\text{SN}} \approx 0.05 \psi_{\text{SF}}$ for a Salpeter IMF, Salpeter 1955) so, if a population of SFGs can be characterised by the distribution of its star-formation rates, its redshift distribution, and its spatial clustering characteristics, the γ-ray luminosity and spatial emission properties of that population can be modelled.

3.1 Star-formation rates

The star-formation rate function, $\psi(\psi_{\text{SF}}, z)$ (SFRF) is the number density of galaxies as a function of their star-formation rate. Its evolution is determined by the underlying history of galaxy assembly, together with gas cooling, feedback (from AGN and stars/stellar end-products) and prior star-formation within galaxies. As such, modelling the SFRF reliably has proven to be a challenging task. To date, various approaches have been adopted, including SAMs (Fontanot et al. 2012; Gruppioni et al. 2015) and hydrodynamic simulations (Davé et al. 2011; Tescari et al. 2014; Katsianis et al. 2017a), with varying degrees of success. Indeed, many previous approaches have been found to yield higher numbers of galaxies at all SFRs compared to observations – a discrepancy often attributed to limitations in the implementation of feedback physics (see Tescari et al. 2014; Katsianis et al. 2017a, for further discussion).

In this work, we adopt a SFRF reference model of Katsianis et al. 2017b, which is obtained from simulations using Virgo Consortium’s Evolution and Assembly of GaLaxies and their Environment (EAGLE) project (Schaye et al. 2015; Crain et al. 2015). We used their reference model (10081504–Ref) as it offered coverage of a large range of SFRs (around $10^{-3}$ to $10^1 M_\odot yr^{-1}$) up to redshifts up to $z \approx 8$. When compared with observationally-determined SFRFs, as discussed in Katsianis et al. 2017b, this was found to under-predict the number of galaxies with SFRs of $10^1$-$10^2 M_\odot yr^{-1}$ at $z > 3$, and the number of objects with SFRs of $10^1$-$10^2 M_\odot yr^{-1}$ at $z < 2$. 6 We note that the Katsianis et al. 2017b model assumes a Chabrier (2003) IMF, but a Salpeter (1955) IMF is adopted in our calculation for the γ-ray luminosity of a galaxy (equation A1) and the luminosity of its dust emission (equation 17). If a Salpeter IMF had been assumed, the resulting SFRs would roughly be a factor of 1.8 higher (Katsianis et al. 2017b). To correct for this discrepancy, we therefore scale the SFRF model accordingly.

Integrating over the SFRF yields the cosmic star-formation rate density (CSFRD),

$$\rho^\star(z) = \int \psi(\psi_{\text{SF}}, z) d\log_{10} \psi_{\text{SF}}$$

where $\psi(\psi_{\text{SF}}, z)$ is the SFRF in units of Mpc$^{-3}$ per decade in $\psi_{\text{SF}}$. Katsianis et al. (2017b) demonstrated that a CSFRD function derived from the baseline 10081504–Ref model was found the exhibit a consistently lower normalisation than that from observation by a factor of 1.5, which may result from the differences with respect to observations discussed above. To account for this, we apply a further multiplicative correction to our model. The CSFRD function

$$\rho^\star(z) = \int \psi(\psi_{\text{SF}}, z) d\log_{10} \psi_{\text{SF}}$$

where $\psi(\psi_{\text{SF}}, z)$ is the SFRF in units of Mpc$^{-3}$ per decade in $\psi_{\text{SF}}$. Katsianis et al. (2017b) demonstrated that a CSFRD function derived from the baseline 10081504–Ref model was found the exhibit a consistently lower normalisation than that from observation by a factor of 1.5, which may result from the differences with respect to observations discussed above. To account for this, we apply a further multiplicative correction to our model. The CSFRD function

Comparison of the Katsianis et al. 2017b reference model is made with SFRFs constructed from UV, IR Hα and radio luminosity functions to facilitate broad SFR and redshift coverage. Observationally-derived SFRFs from Mauch & Saldier (2007); Reddy et al. (2008); Gilbank et al. (2010); Rodighiero et al. (2010); Karim et al. (2011); Ly et al. (2011); Robotham et al. (2011); Gruppioni et al. (2013); Magnelli et al. (2013); Patel et al. (2013); Sobral et al. (2013); Bouwens et al. (2015); Alavi et al. (2016); Marchetti et al. (2016); Parsa et al. (2016) as well as from compiled data (Madau & Dickinson 2014), are used. Additionally, comparison is made with SFRFs from Smit et al. (2012); Duncan et al. (2014); Katsianis et al. (2017a).
derived from the 100N1504--Ref model was otherwise largely consistent with observations (Gilbank et al. 2010; Karim et al. 2011; Robotham et al. 2011; Sobral et al. 2013; Madau & Dickinson 2014; Bouwens et al. 2015) with the exception of that obtained from IR data (Rodighiero et al. 2010; Madau & Dickinson 2014). This was considered to be due to assumed dust corrections in computing UV luminosities, incomplete UV luminosity functions or possible over-estimations of the SFR from IR data (Katsianis et al. 2017b). It was further shown that those SFGs with high star-formation rates, between 10 and 100 M$_{\odot}$ yr$^{-1}$, exhibit the strongest redshift dependence, peaking sharply at $z \sim 2$, while less vibrantly star-forming galaxies show a weaker evolution in their contribution to the CSFRD (this is also in tension with IR studies, e.g. Magnelli et al. 2013, which do not find such a strongly peaked evolution of highly star-forming galaxies). These intensively star-forming galaxies presumably represent the most important SFG sub-class contribution to the EGB, which should also reflect this strongly peaked evolutionary history. It will be shown in the following sections (3.2 and 3.3) that this would imprint a distinctive spatial signature into the EGB.

3.2 Clustering and bias

In the hierarchical model of structure formation, spatial clustering of galaxies is primarily determined by the distribution of dark matter in the Universe. Dark matter haloes form from the gravitational collapse of primordial Gaussian density perturbations, with their density (which are influenced by the gravitational potential, and hence mass, of the halo – see Silk & Wyse 1993; Sutherland & Dopita 1993). The subsequent evolution of galaxies through cosmic time experiencing accretion of new gas from the cosmic web, feedback and mergers yields the properties of populations of galaxies at high redshifts (Jose et al. 2013; Harikane et al. 2018) and, eventually in the present Universe (e.g. Press & Schechter 1974; Lacey & Cole 1993; Sheth & Tormen 1999; Behroozi et al. 2013) and so form biased tracers of the underlying dark matter distribution of the Universe at different epochs (e.g. Kaiser 1984; Cooray & Sheth 2002; Mo et al. 2010).

The bias of galaxy population clustering compared to that of dark matter is typically studied observationally from their spatial distribution, with various sources classes having been found to exhibit different clustering properties (e.g. see Hale et al. 2018, which finds a different clustering bias for AGNs and SFGs against dark matter, with AGNs typically exhibiting greater clustering strength). We define the effective clustering bias factor of SFGs compared to dark matter using the relation $b_{SFG}(k, z) \equiv b_{SFG} \langle P_{SFG}(k, z) \rangle P_{DM}(k, z)$, where $P_{SFG}(k, z)$ is the power spectrum of SFGs, and $P_{DM}(k, z)$ is the power spectrum of linear dark matter density fluctuations. We calculate $P_{DM}(k, z)$ using the transfer function approximation of Eisenstein & Hu (1999), which is shown to be accurate to within 5%.

The SFG population bias factor, $b_{SFG}$ may be calculated from the ratio of galaxy to dark matter correlation functions, i.e:

$$b_{SFG}^2(z) = \frac{\xi_{g}(r, z)}{\xi_{DM}(r, z)} = \left( \frac{\delta_{0}(z)}{\sigma_{g}^{8}(r, z)} \right) \frac{J_{2}}{8} \quad (27)$$

(e.g. Kaiser 1984; Bardeen et al. 1986; Lindsay et al. 2014), where the matter fluctuation amplitude $\sigma_{g} = 0.811 \pm 0.006$ (Planck Collaboration et al. 2020), and $\delta_{0}(z) = g(\delta_{g}(z)/g_{0})$, with $g(z)$ as the growth factor at redshift $z$ and $g_{0} = g(z = 0)$ (e.g. Carroll et al. 1992). Additionally, $J_{2} = 72/([3 - \ell][4 - \ell][6 - \ell]^{2})$, and $\delta_{0}(z) = r_{0}^{2}(1 + z)^{p}$ with $p = 1 - (3 + w)/\ell$ (Lindsay et al. 2014). Here, the choice of the parameter $w$ reflects the clustering model adopted. In this demonstrative model we consider only linear clustering (Overzier et al. 2003) where clustering growth is set by linear perturbation theory and $w = -1$. We leave the investigation of alternative clustering growth models to future work – for example, stable clustering (where clusters have a fixed physical size and $w = 0$), co-moving clustering (where clusters have fixed co-moving size and $w = -3$) and decaying clustering (which implies a rapid clustering decay) are also considered in the literature (Overzier et al. 2003; Kim et al. 2011; Elvity et al. 2012). The remaining parameters in equation 27 are the power-law slope of the two-point correlation function of galaxies, $\ell$, and the galaxy clustering length $r_{0}^{2}$. Both of these may be estimated empirically for SFGs, and we adopt the best-fit values of Hale et al. (2018): $\ell = 1.8$ and $r_{0}^{2} = 6.1$ Mpc $h^{-1}$. These were computed from radio-selected SFGs in the COSMOS field using deep Karl G. Jansky Very Large Array (VLA) data at 3 GHz, reaching redshifts as high as $z = 5$, thus covering our range of interest ($z \leq 3$). The resulting bias factor from these parameter choices is higher than those computed for SFGs at other wavelengths (e.g. Gilli et al. 2007; Starikova et al. 2012; Magliocchetti et al. 2013), but this is attributed to the greater extent of the redshift distribution of the sources.

3.3 Development of EGB anisotropies

The SFG power spectrum $P_{SFG}(k, z)$ would imprint a signature in the EGB, even though individual contributing sources would not typically be resolved. The distribution of spatial scales of this signature would depend on redshift $z$, being specified by $P_{SFG}(k, z)$, and the strength of the contribution from a shell in redshift would correspond to the relative $\gamma$-ray luminosity of the source population at that epoch, as set by the SFG redshift distribution (see schematic in Figure 2). This could be measured from $\gamma$-ray background observations using the auto-correlation function (equation D1), from which a clustering term $C_{E}^{g}$ and an isotropic Poisson noise term (an auto-correlation term) $C_{E}^{P}$ can be decomposed from the Fourier Transform (see Appendix D for details). These may be written as

$$C_{E}^{g}(E_{\gamma}) = \int_{0}^{\infty} \frac{dV_{E}}{dV} \left( \frac{\delta}{P_{\gamma}^{2} + 1} \right) \left( \frac{dF_{E}(E_{\gamma}, z)}{dE_{\gamma}} \right)^{2}, \quad (28)$$

7 We calculate this using the formula presented in Hamilton 2001, using the public code provided at: https://jila.colorado.edu/~ajsh/growl/.

8 We note that Magliocchetti et al. 2017 also provided values for these parameters for a radio-selected sample of SFGs at 1.4 GHz (with radio fluxes above 0.15 mJy) up to $z \sim 3$. However, the number of data points in their analysis is much fewer than in Hale et al. (2018), leading to our preference to use the best-fit values of the later study.

MNRAS 000, 1–21 (2020)
Schematic to illustrate the redshift evolution of the power spectrum of the source population as a function of multipole $\ell$. The peak would correspond to the multipole the dominant signal is imprinted at. Here, $z_1 < z_2 < z_3$.

![Figure 2](image)

Figure 2: Schematic to illustrate the redshift evolution of the power spectrum of the source population as a function of multipole $\ell$. The peak would correspond to the multipole the dominant signal is imprinted at. Here, $z_1 < z_2 < z_3$.

The approach of Peretti et al. (2020), line 7, is broadly consistent with the prediction of this work, with some deviations at lower energies and a smaller cascade contribution (line 9). The low-energy difference is likely accounted for by the additional physics included in the spectral model of Peretti et al. (2020) that would boost the low-energy $\gamma$-ray flux compared to this work (for example, their inclusion of inverse-Compton and bremsstrahlung emission may become relatively important in lower star-formation rate sources, where pion-decay $\gamma$-ray emission would be less dominant). The differences in the cascade prediction between this work and that of Peretti et al. (2020) would presumably arise from their delta-function approximation of the EBL radiation field, compared our use of the Inoue et al. 2013a EBL model. Given the current uncertainties in EBL models, it can be reasonably argued that both approaches to the cascade emission are equally valid, and that future estimations of the cascade contribution will improve as observational constraints on the EBL are tightened.

4 RESULTS AND DISCUSSION

4.1 EGB spectrum

The EGB spectrum between 0.1 and 50 GeV predicted at $z = 0$ by our fiducial model, which adopts a characteristic SFG nucleus of $R = 0.1$ kpc, a CR spectral index $\Gamma' = 2.1$ and a maximum CR energy $E_{\text{max}} = 50$ PeV, is shown in Figure 3. Here, both the total contribution to the diffuse EGB from SFGs (line 1), and that arising from the cascaded SFG emission (line 2) are included. For comparison, the contribution from resolved and unresolved blazars is shown (band 3, denoting the range of 3 models presented in Ajello et al. 2015 – however, these do not include a cascade flux component), together with the total observed diffuse EGB spectrum using 50 months of Fermi-LAT data (taken from Ajello et al. 2015, with original data from Ackermann et al. 2015). For reference, the contribution from all AGN presented in Makiya et al. (2011) is also shown (line 4). Makiya et al. (2011) also compute the contribution from SFGs (line 5), which we find to be substantially lower than many other literature models. It can be seen that our fiducial model is in agreement with the observational constraints given by the contribution to the EGB from resolved and unresolved blazars, however the predicted SFG contribution comes close to saturating the diffuse EGB at higher energies, above a few 10s GeV (but remains compatible with observational limits). This behaviour is also evident in some other models, e.g. Peretti et al. 2020 (line 7).

In Figure 3, the substantial variation in predictions made by other models is clear. Here, we draw comparison between our fiducial model and those in the literature which consider a contribution specifically from SFGs. We find our approach yields a $z = 0$ EGB intensity that is much higher than the the SAM-based method considered by Makiya et al. 2011 (also that of Lamastra et al. 2017, which falls substantially lower even than the Makiya et al. 2011 prediction, and is not shown in Figure 3), which is exceeded by as much as an order of magnitude at energies above $\sim$ 10 GeV. Both the Makiya et al. 2011 and Lamastra et al. 2017 SAM-based models are strongly dependent on the source population properties, redshift distributions and $\gamma$-ray emission models adopted, all of which differ compared to equivalent model components adopted in this work.

By contrast, the SFG contribution intensity computed by Chakraborty & Fields (2013), line 6, is substantially higher than our prediction. It also exceeds predictions by other models up to energies of $\sim$ 3 GeV, as shown. It is even comparable to the all blazar contribution of Ajello et al. (2015) below $\sim$ 0.6 GeV. However, its steeper power-law in energy, resulting from the steeper assumed CR proton spectrum within the source population, causes the Chakraborty & Fields (2013) model to have fallen far below the prediction of this work by 50 GeV.

The approach of Peretti et al. (2020), line 7, is broadly consistent with the prediction of this work, with some deviations at lower energies and a smaller cascade contribution (line 9). The low-energy difference is likely accounted for by the additional physics included in the spectral model of Peretti et al. (2020) that would boost the low-energy $\gamma$-ray flux compared to this work (for example, their inclusion of inverse-Compton and bremsstrahlung emission may become relatively important in lower star-formation rate sources, where pion-decay $\gamma$-ray emission would be less dominant). The differences in the cascade prediction between this work and that of Peretti et al. (2020) would presumably arise from their delta-function approximation of the EBL radiation field, compared our use of the Inoue et al. 2013a EBL model. Given the current uncertainties in EBL models, it can be reasonably argued that both approaches to the cascade emission are equally valid, and that future estimations of the cascade contribution will improve as observational constraints on the EBL are tightened.

4.2 EGB anisotropy signatures

We directly compute the EGB intensity fluctuation angular power spectrum at $z = 0$ arising from our model SFG population. This uses the computational method outlined in Appendix E to solve equations 28 and 29. Large numbers of photons are needed to compute high-resolution spectral statistics from data. Typically, $\gamma$-ray data analysis methods would bin events according to photon energy, to improve signal-to-noise ratios within an energy band and to reduce the requirement on photon numbers in a small energy range. We therefore compute our expected anisotropy signatures in broad energy bins to reflect this. Figure 4 shows the EGB anisotropy signature computed for our fiducial model, integrated over the energy band $E_\gamma = (1 - 10)$ GeV. Uncertainties from the empirical dust relation of equation 18 were propagated, but found to be negligible. While the total EGB anisotropy signature is plotted in this case, the clustering contribution (cf. equation D2) exceeds the Poisson component by around 3 orders of magnitude – consistent with the expectation that the Poisson (statistical noise) contribution from a source population comprised of a large number of unresolved faint galaxies would be relatively low.
4.2.1 Energy bands

The intensity of the EGB varies with energy (cf. Figure 3). As the cosmological attenuation of γ-rays is also energy-dependent, with stronger flux suppression arising at higher energies (e.g. Gilmore et al. 2009; Inoue et al. 2013a), the EGB anisotropy would differ according to the choice of energy band. Figure 5 demonstrates that such differences are almost negligible, when comparing the EGB angular power spectrum in four bands, (0.1-1) GeV, (1.0-10) GeV, (10-20) GeV and (20-30) GeV for the fiducial model. The upper panel shows the main difference between these four energy bands follows simply from the EGB energy spectrum (Figure 5).

4.2.2 Model parameters

The three fixed parameters in the fiducial model are $R$, $\Gamma$ and $E_{\text{max}}$. However, some variation of their values would be expected throughout a real SFG source population, with implications for the EGB intensity and anisotropy. The radius $R$ of a SFG nuclear region could vary substantially between galaxies. For example, among starburst galaxies in the local Universe, it is found to differ by a factor of a few – in NGC 253, $R \approx 0.1$ kpc (Weaver et al. 2002), while

\begin{align*}
E_y &= (1 - 10) \text{GeV} \\
C_\ell^P &= 1.22 e^{-12}
\end{align*}
for M82, \( R \approx 0.25 \text{ kpc} \) (de Grijs 2001). Moreover, in models and simulation work, compact galaxies are found to be common at high-redshift (e.g. Furlong et al. 2017), which would imply a redshift-dependence in \( R \) for realistic SFG source distribution models. Such variations would have discernible effects on the EGB intensity and anisotropy. The impact of alternative choices of \( R \), with the value increased and decreased by a factor of 2 compared to the fiducial choice of 0.1 kpc are shown in Figure 6. This demonstrates the EGB intensity is directly affected by the value of \( R \) set in the source population, with higher intensities developing for a larger characteristic choice of \( R \). This effect can be understood from the spatial spread of photons through a SFG nucleus when \( R_{\text{SFG}} \) is fixed. Increasing \( R \) would increase the volume of the SFG nucleus, and decrease the photon density in the stellar and dust radiation fields that attenuate \( \gamma \)-rays. More \( \gamma \)-rays would then escape from their source galaxy, contributing more photons to the EGB. Anisotropies are unaffected in this case, as \( R \) is adjusted independently of redshift. If a more physical redshift-dependent treatment of \( R \) were adopted, an anisotropic signature would presumably emerge in the EGB. However, the necessary detailed modelling of appropriate redshift-size relations for SFG populations falls beyond the scope of this study, and is left to future dedicated work.

A similar comparison for variation of \( \Gamma \) is shown in Figure 7, where values of \( \Gamma = 1.9 \) and 2.3 are considered alongside the fiducial choice. These represent a less-steep (steeper) internal proton spectrum in the SFG population (respectively), as may arise from a younger (older) CR spectrum, or due to variations in accelerator geometries/configurations or CR acceleration physics, and reflects the range of values determined from observational analyses of nearby \( \gamma \)-ray emitting SFGs (Ajello et al. 2020). The impact of this variation is a change in \( \gamma \)-ray flux (and hence \( C_{\ell} \) normalisation), as shown in the upper panel, with an increased EGB intensity for a steeper choice of CR index. As the \( \gamma \)-ray emission spectrum from the SFG closely reflects the hadronic CR spectrum, a steeper CR spectral index yields more power in the \( \gamma \)-ray energy spectrum at lower energies. From Figure 1, it can be seen that the strongest attenuation from the source galaxy is felt by higher energy \( \gamma \)-rays, so the fraction attenuated within SFGs is reduced for steeper CR spectral indices. The lower panel of Figure 7 reveals the shape of the EGB anisotropy power spectrum is also influenced by the choice of \( \Gamma \), where a steeper CR spectrum yields a noticeably sharper EGB angular power peak, while a softer CR spectrum produces a broader peak. This effect follows from the energy dependence of \( \gamma \)-ray attenuation in the EBL; despite the internal attenuation, a less steep CR source spectrum would ultimately still produce a higher fraction of high-energy \( \gamma \)-rays. These are attenuated more readily, and fewer photons from distant sources survive to \( z = 0 \), even when considering the cascade process. The fraction of flux contributed by SFGs at large distances (which would imprint signatures on larger angular scales) is therefore reduced for less steep CR spectra, effectively suppressing EGB angular power, particularly on larger scales. Recent work has considered the possibility of blended spectral indices within SFGs (Ambroso et al. 2021). The results here would imply these would have a non-trivial impact on the EGB anisotropy, and should be explored further in future studies.

The upper limit of the CR spectrum in SFGs is determined by acceleration mechanisms and the detailed configuration of the accelerators (e.g. for discussion, see Peretti et al. 2020), and the exact value that should be adopted in any given environment remains unsettled. However, we find this is not of particular consequence to our results. Figure 8 considers alternative choices of \( E_{\text{max}} \), which shows a limited effect on the EGB intensity. Only a small intensity boost is seen if a lower maximum cut-off is adopted, or a proportionally small decrease arises if a higher cut-off is instead chosen. This can be predominantly accounted for by the adjustment in the spectral normalisation for different choices of \( E_{\text{max}} \) (see equation A7), rather than any physical process. The EGB anisotropy is not dependent on the exact choice of \( E_{\text{max}} \).

### 4.2.3 Alternative redshift evolution scenarios

Our fiducial model adopts the galaxy population model of Katsianis et al. 2017b, which yields a redshift distribution of cosmic star-formation broadly compatible with Madau & Dickinson (2014), where the peak of cosmic star-formation arises at \( z \sim 2 \). However, this may not fully reflect the diverse redshift distributions of various classes of SFGs (e.g. the distribution of sub-mm galaxy samples in Simpson et al. 2014 compared to that of the luminous sub-mm sources in Koprowski et al. 2014 or dusty star-forming galaxies in Strandet et al. 2016), which are not guaranteed to follow the global mean cosmic star-formation history (CSFH) of the Universe. We crudely demonstrate the level of impact alternative CSFHs would have on the EGB in Figure 9, where we modify our fiducial distribution derived from Katsianis et al. 2017b by simply adjusting its redshift distribution by \( \pm 0.5 \), thus creating an ‘early’ CSFH model, and a ‘delayed’ CSFH model. The main impact of this is on the EGB intensity, which is reduced for the earlier CSFH.
model, or increased for the later one (see Figure 9). This follows largely from our crude adjustment, in that more stars would form in the ‘early’ CSFH scenario (and conversely, fewer in the ‘delayed’ CSFH). However, more subtle effects emerge in the EGB angular power spectrum (Figure 9, lower panel). It is not intuitive that the spectral shape is broadened both in the ‘early’ and ‘delayed’ CSFH scenarios compared to the fiducial model, with a slightly greater skew towards more power at larger $\ell$s (smaller scales). These can both be understood from the interplay between the redshift distribution of sources in a spherical volume, and the attenuation of $\gamma$-rays in EBL radiation fields: in the ‘early’ CSFH model, there are more sources at higher redshift (imprinting EGB signatures on larger angular scales). However, the greater distance to these sources means a greater degree of $\gamma$-ray attenuation in the intervening EBL, so their contribution (per source) to the $z = 0$ EGB would be relatively weak. The is partially compensated by the larger number of sources contained within the volume to a higher redshift, thus broadening the EGB anisotropy signature slightly more over a wider range of scales compared to the fiducial model - i.e. making it less strongly peaked. The converse is true for the ‘delayed’ CSFH model, but the effect is broadly the same due to the EBL attenuation and source distribution acting antagonistically.

While these crude variations in CSFH offer little physical insight into the astrophysics of SFG populations, they do illustrate that signatures imprinted by SFGs are influenced by their redshift distribution, and that both its intensity and anisotropy encode information about this. EGB anisotropies particularly offer potential as a diagnostic tool to distinguish between different redshift distributions of source populations and, hence, offer scope as a probe the evolutionary histories of population classes of SFGs in which CR activity is important. However, we have shown that these signatures can be subtle, and must be carefully modelled and understood before they can be reliably used to probe and interpret CR activity within source distributions over redshift.

Figure 6. EGB intensity fluctuation angular power spectrum $C_\ell$ shown against multipole $\ell$, for alternative choices of the characteristic starburst nucleus size of the source population from $R = 0.1$ kpc (fiducial value) to $R = 0.2$ kpc and 0.05 kpc. While there is a strong impact on the EGB intensity for difference choices of $R$, the angular power spectrum shape is not affected. Results are shown for the energy band $E_\gamma = (1–10)$ GeV.

Figure 7. EGB intensity fluctuation angular power spectrum $C_\ell$ shown against multipole $\ell$, for alternative choices of the spectral index of the hadronic CRs, $\Gamma$, in the source population. Top and bottom panels as per Figure 5. Results are shown for the $\gamma$-ray energy band $E_\gamma = (1–10)$ GeV.

Figure 8. EGB intensity fluctuation angular power spectrum $C_\ell$ shown against multipole $\ell$, for alternative choices of maximum CR energy, $E_{\text{max}}$. The impact on the intensity shown here results from the dependence of the CR spectral normalisation on $E_{\text{max}}$ (see equation A7), and is not physical. There is no impact on the shape of the EGB anisotropy. Results are shown for the energy band $E_\gamma = (1–10)$ GeV.
The star-forming galaxy contribution to the EGB

4.3 Observational prospects

4.3.1 Statistical error

The projected statistical 1-σ error in an extracted measurement of \( C_\ell \) is given by

\[
\delta C^2_\ell = \frac{2C^2_\ell}{(2\ell + 1)\Delta \ell \Delta f_{\text{sky}}},
\]

(Ando et al. 2007a,b), where \( \Delta \ell \) is the bin size in multipole \( \ell \), \( f_{\text{sky}} \) is the fraction of sky covered by the relevant γ-ray survey. We find this error dominates over all uncertainties built into our model, and would be the primary limitation in resolving EGB signatures. We show this projected statistical error in \( C_\ell \) for 40 equal bins in log \( \ell \) space for \( f_{\text{sky}} = 0.25 \) in black (this is indicative of the sky coverage anticipated as part of CTA’s Extra-galactic Survey Key Science Project – see CTA Consortium 2019 for details), and \( f_{\text{sky}} = 1 \) in red (reflective of the full-sky coverage of Fermi-LAT) in Figure 10. It is evident that low multi-poles, or large scale anisotropies are most affected by statistical fluctuations. At intermediate and small scales, the statistical error is greatly reduced (smaller scale anisotropies are computed by splitting the sky into a larger number of regions, thus reducing statistical variations), with good prospects for signal extraction.

4.3.2 Integration time

The integration time required to detect EGB anisotropies can be estimated by comparing the γ-ray background intensity with instrument sensitivity. To demonstrate the prospects for detecting a signal, we consider the sensitivities of the current Fermi-LAT observatory\(^9\) and those estimated for the up-coming CTA,\(^10\) at 50 GeV, where reasonable comparison may be made between the two instruments, and compute the corresponding integration time for our fiducial model in each case. The mean intensity of the EGB at 50 GeV was found to be \( 9.9 \times 10^{-8} \) GeV cm\(^{-2}\) s\(^{-1}\) sr\(^{-1}\) in our fiducial model (see Figure 3), with anisotropic variations leading to minimum intensities on the scales of interest reaching around 5% of this value. As such, we argue it would be necessary to detect EGB intensities as low as \( 5.0 \times 10^{-9} \) GeV cm\(^{-2}\) s\(^{-1}\) sr\(^{-1}\) in our fiducial model to clearly recover anisotropy signatures. This corresponds to a flux threshold of \( 7.9 \times 10^{-12} \) erg cm\(^{-2}\) s\(^{-1}\), which would be detectable with Fermi-LAT (at a 5\(\sigma\) level and with at least 10 counts per bin) after around 5.5 years of observation. This estimate assumes uniform sky exposure. In reality, the exposure of Fermi-LAT varies by a factor of 0.57 (e.g. Nolan et al. 2012), and so our estimate should be correspondingly increased to at least 10 years for a signature to be observed. The 10-year Pass 8 release of LAT data is therefore already reaching sufficient (or near-sufficient) exposure on many scales to detect anisotropies attributed to SFG populations, and accordingly initial detections of EGB anisotropy signatures from SFG populations are emerging (Fornasa et al. 2016; Ackermann et al.

\(^9\) Fermi-LAT top-level Pass 8 performance information is available online, see https://www.slac.stanford.edu/exp/glast/groups/canda/lat_Performance.htm. Instrument response functions are based on the new event analysis and selection criteria described by Atwood et al. 2013.

\(^{10}\) CTA instrument response functions are provided by the CTA Consortium and Observatory, see http://www.cta-observatory.org/science/cta-performance/ (version prod3b-v2) for more details.
These will improve over time, as Fermi-LAT integration time continues to increase. The projected CTA integration time to detect the same intensity would be around 0.5 hours (this is approximated from the estimated sensitivity of the CTA-North array at a 70-degree elevation angle). Given that the proposed CTA extragalactic Survey Key Science Project would cover around 25% of the extra-galactic γ-ray sky with a uniform integration time of 1.11 hours (CTA Consortium 2019), EGB intensities around 2 times fainter could be reached, depending on the final array configuration and observational strategy adopted. This greatly improves the potential for resolving γ-ray signatures, and opens the prospect for much more detailed signature extraction.

5 SUMMARY & CONCLUSIONS

This work has shown how signatures are imprinted into the EGB and demonstrated how their contribution may be characterised using a small number of physically-motivated parameters. Moreover, it has outlined the relevant EGB statistics that can be used to probe the evolution of the underlying source populations, and has provided a proof-of-concept example by showing the EGB anisotropic signature expected to arise from a population of SFGs. This signature is dominated by the contribution from SFGs around the so-called ‘high noon’ of star-formation at redshifts $z \sim 2-3$, where physical conditions and processes in galaxies differ dramatically from those in the local Universe. The interactions of CRs, their associated production of particles and radiation, and their deposition of momentum during this epoch become important factors in controlling the evolution of galaxies and producing energetic cosmic backgrounds. The EGB offers scope to probe these interactions in a direct way, and analysis of patterns within the EGB offer potential to advance our understanding of critical aspects of CR interactions in and around SFGs, in particular during the cosmic noon.

We have further shown that different sub-populations of SFGs could be resolved by a careful analysis of the EGB intensity and angular power spectrum, once appropriate physical models and signal extraction techniques are developed, and that Fermi-LAT will soon reach sufficient integration times for signatures imprinted by SFGs to be extracted. This will be substantially improved by up-coming facilities, e.g. CTA, which will offer far greater sensitivities and will be able to resolve SFG source populations in even more detail. We have also demonstrated that intergalactic magnetic fields can distort imprinted EGB signatures, but the magnitude and structure of this distortion is currently unclear. It is essential that this is explored carefully in future work to ensure that physical interpretations of EGB anisotropies can be reliably made.

In the coming decade, a wealth of new EGB data will become available to the γ-ray community, with current and up-coming instruments offering unprecedented sensitivities and resolution. There is great potential to use this data to infer new information about cosmic star-formation, intergalactic magnetic fields, SFGs and their properties, and AGN – if appropriate models for the detailed signatures these imprint in the EGB are available. However, theoretical and methodological frameworks must first be urgently developed to ensure efforts in the community are able to make optimal use of this up-coming data and the opportunities it presents.

ACKNOWLEDGEMENTS

This work used high-performance computing facilities operated by the Center for Informatics and Computation in Astronomy (CICA) at National Tsing Hua University (NTHU). This equipment was funded by the Ministry of Education of Taiwan and the Ministry of Science and Technology of Taiwan. We are also grateful to the National Center for High-performance Computing (Taiwan) for computer time and facilities. ERO is supported by the Ministry of Education of Taiwan at CICA, NTHU. His visits to Kavli IPMU were hosted by KGL, and supported by a travel grant of University College London’s Mullard Space Science Laboratory (UCL/MSSL), and the Ministry of Science and Technology of Taiwan. KGL acknowledges support from JSPS KAKENHI Grants JP18H05868 and JP19K14755. This research has made use of the CTA instrument response functions provided by the CTA Consortium and Observatory, see http://www.cta-observatory.org/science/cta-performance/ (version prod3b-v2) for details. ERO thanks Prof. John Silverman (IPMU) and Prof. Yoshiyuki Inoue (Osaka) for helpful discussions about star-forming galaxies.  

MNRAS 000, 1–21 (2020)
Figure 11. Schematic to illustrate the development of EGB signatures from an evolving unresolved source population. Two redshift slices at \( z_1 \) and \( z_2 \), where \( z_2 > z_1 \) are shown, in addition to the \( \gamma \)-ray sky (grid) at \( z = 0 \). The separation scale \( k_2 \) at \( z_2 \) imprints a signature of intensity \( I_2 \) with characteristic separation \( \ell_2 \) on the EGB formed at \( z = 0 \). At \( z_1 \), the separation scale \( k_1 \) imprints a signature of intensity \( I_1 \) with characteristic separation \( \ell_1 \), where \( \ell_1 < \ell_2 \). The signatures are reprocessed by the EBL as it propagates through cosmological space, with red colouring indicating an increased proportion of secondary cascaded \( \gamma \)-rays. The cascade fraction is greater for a beam originating at higher redshift, but the reprocessing does not affect the signal anisotropy pattern directly unless magnetic fields are invoked.

and the cosmological propagation and attenuation of \( \gamma \)-rays, Prof. Masahiro Teshima for facilitating visits to the Institute of Cosmic Ray Research, University of Tokyo, where useful discussions took place, and Prof. John S. Gallagher III (UW-Madison) and Prof. Tomotsugu Goto (NTHU) for comments about the implications of this work. The authors thank Prof. Kinwah Wu (UCL/MSSL) for discussions that helped to inform and inspire the early stages of this work, and the anonymous referee for their detailed review which substantially improved the manuscript. This research used NASA's Astrophysics Data Systems.

**DATA AVAILABILITY**

No new data were generated or analysed in support of this research.

**REFERENCES**

Abdo A. A., et al., 2010a, ApJ, 709, L152
Abdo A. A., et al., 2010b, ApJ, 720, 435
Abeysekara A. U., et al., 2019, ApJ, 885, 150
Acciari V. A., et al., 2019, MNRAS, 486, 4233
Acero F., et al., 2009, Science, 326, 1080
Ackermann M., et al., 2012, ApJ, 755, 164
Ackermann M., et al., 2015, ApJ, 799, 86
Ackermann M., et al., 2018, Phys. Rev. Lett., 121, 241101
Aharonian F. A., Coppi P. S., Voelk H. J., 1994, ApJ, 423, L5
Ahlers M., Salvado J., 2011, Phys. Rev. D, 84, 085019
Ajello M., et al., 2012, ApJ, 751, 108
Ajello M., et al., 2014, ApJ, 780, 73
Ajello M., et al., 2015, ApJ, 800, L27
Ajello M., Di Mauro M., Paliva V. S., Garrappa S., 2020, ApJ, 894, 88
Ajiki M., et al., 2002, ApJ, 576, L25
Alavi A., et al., 2016, ApJ, 832, 56
Allard D., Parizot E., Olinto A., 2007, Astroparticle Physics, 27, 61
Almeida S. P., 1968, Phys. Rev., 174, 1638
Ambrosone A., Chianese M., Fiorillo D. F. G., Marinelli A., Miele G., Pisanti O., 2021, MNRAS, 503, 4032
Ando S., Komatsu E., Narumoto T., Totani T., 2007a, Phys. Rev. D, 75, 063519
Ando S., Komatsu E., Narumoto T., Totani T., 2007b, MNRAS, 376, 1635
Andrews S. K., Driver S. P., Davies L. J. M., Lagos C. d. P., Robotham A. S. G., 2018, MNRAS, 474, 898
Atwood W., et al., 2013, arXiv e-prints, p. arXiv:1303.3514
Axford W. I., Leer E., Skadron G., 1977, in International Cosmic Ray Conference, p. 132
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bassett R., et al., 2017, MNRAS, 467, 239
Beck A. M., Lesch H., Dolag K., Kortarba H., Geng A., Stasyszyn F. A., 2012, MNRAS, 422, 2152
Berezhetskii V. B., Lifshits E. M., Pitaevskii L. P., 1980, Moscow Izdatel Nauka Teoreticheskai Fiizika
Berezinsky V. S., Bulanov S. V., Dogiel V. A., Ptuskin V. S., 1990, Astroparticle Physics of cosmic rays, Amsterdam: North-Holland
Berezinsky V., Gazizov A., Kachelrieß M., Ostapchenko S., 2011, Physics Letters B, 695, 13
Bergin E. A., Tafalla M., 2007, ARAA, 45, 339
Bergvall N., Marquart T., Way M. J., Blomqvist A., Holst E., Östlin G., Zackrisson E., 2016, A&A, 587, A72
Bernet M. L., Miniati F., Lilly S. J., Kronberg P. P., Duseaegues-Zavsky M., 2008, Nature, 454, 302
Berta S., et al., 2011, A&A, 532, A49
Bertone G., Hooper D., Silk J., 2005, Phys. Rep., 405, 279
Béthermin M., et al., 2012, A&A, 542, A58
Béthermin M., et al., 2015, A&A, 573, A113
Bhattacharya D., Sreekumar P., 2009, Research in Astronomy and Astrophysics, 9, 509
Blandford R. D., Ostriker J. P., 1978, ApJ, 221, L29
Blasi P., 2011, in Giani S., Leroy C., Rancoita P. G., eds, Cosmic Rays for Particle and Astroparticle Physics, pp 493–506 (arXiv:1012.5005), doi:10.1142/9789814329033_0061
Blattnig S. R., Swaminathan S., Kruger A. T., Ngom M., Norbury J. W., Tripathi R. K., 2000, Technical report, Parameterized Cross Sections for Pion Production in Proton-Proton Collisions
Blumenthal G. R., Gould R. J., 1970, Reviews of Modern Physics, 42, 237
APPENDIX A: COSMIC RAY PROTON DENSITY

Strong star-forming activity gives rise to frequent SN events and remnants shortly after the onset of star-formation. Such environments can accelerate particles to relativistic energies through, e.g., Fermi (1949) acceleration in diffuse magnetised shocks (Axford et al. 1977; Krymskii 1977; Blandford & Ostriker 1978; Bell 1978a,b) – see Blasi (2011) for further discussion. The energy injected into the relativistic particles (i.e., CRs) is governed by the total energy provided in the SN events, and the efficiency of acceleration. To describe their total luminosity, we adopt the relation

\[ L_{\text{CR},\text{eff}} = \varepsilon f_v E_{\text{SN}} \rho_{\text{SN}} = \alpha \left[ \frac{f_v E_{\text{SN}} \rho_{\text{SF}}}{M_{\text{SN}}} \right]. \]  

(\text{A1})

(Owen et al. 2018), where \( E_{\text{SN}} \) is the SN event total energy (around \( 10^{51} \) erg for core-collapse Type II P SNe expected to dominate in these highly star-forming systems), \( \varepsilon \) is the CR acceleration efficiency (0.1 is set as a conservative value in this work, although some variation could be justified – see Fields et al. 2001; Strong et al. 2010; Lemoine-Goumard et al. 2012; Caprioli 2012; Dermer & Pacaud 2013; Morlino & Caprioli 2012; Wang & Fields 2018 for indications on the possible range) while \( f_v = 0.01 \) is the fraction of energy retained by the SN event after neutrino losses (although we adopt a value of 1% here, the exact choice would depend on the SN type and environment with fractions as low as 0.1% being plausible – for discussion, see Iwamoto & Kunugise 2006; Smartt 2009; Janka 2012). \( \alpha \) indicates the fraction of stars which evolve to produce a core-collapse SN event, and this is governed by the initial stellar mass function (IMF) and its upper cut-off. For a Salpeter IMF with index \( Y = 2.35 \) between 1 M\(_{\odot}\) and an upper cut-off mass for stars able to ultimately produce a SN event of \( M_{\text{SN}} = 50 M_{\odot} \) (see, e.g., Fryer 1999; Heger et al. 2003), this fraction takes a value of \( \alpha = 0.05 \), which we fix for all galaxies in the present study (this was also adopted in earlier work, e.g., Owen et al. 2018), assuming no variation in the stellar mass function between galaxies. While future work may consider alternative mass function models and upper values for the mass cut-off, their bearing on the CR flux is relatively weak. \( \rho_{\text{SN}} \) and \( \rho_{\text{SF}} \) are the SN event rate and star-formation rate of a galaxy, respectively. Either may be used to parameterise the CR flux normalisation, with the former being more closely related to the energy-injection rate (and hence of closer bearing to the physical model), while the latter may be considered a closer relation to observational quantities. Hereafter, we specify the CR luminosity of galaxies in terms of \( \rho_{\text{SF}} \) and leave all other parameters in equation A1 fixed at the stated values.

The CR proton density at some location \( r \) within a galaxy may be modelled as the superposition of the contributions from an ensemble of \( N \) continuously-injecting discrete CR sources located at the points described by the position vectors \( r_i \),

\[ n_p(\gamma_p, r) \, d\gamma_p = \sum_{i=1}^{N} f_{adv} Q_p(\gamma_p, r) Y A(\gamma_p, r) \, d\gamma_p, \]  

(A2)

(Owen et al. 2019b), once a galaxy has settled into a steady-state with a roughly constant star-formation rate and saturated magnetic field. This is typically reached a few tens of Myr after the onset of star-formation for SFGs, with the magnetic saturation time being inversely proportional to \( \rho_{\text{SF}} \) (Schober et al. 2013). D is the energy-dependent CR diffusion coefficient (see section A1 for details), which takes the parametric form

\[ D(\gamma_p) = D_0 \left( \frac{r_L(\gamma_p, [B])}{r_{L,0}} \right)^5, \]  

(A3)

where \([B] = |B|\) is the characteristic interstellar magnetic field strength. The normalisation value \( D_0 = 3.0 \times 10^{28} \) cm\(^2\) s\(^{-1}\) is based on empirical measurements of the diffusion of CRs in the interstellar medium (ISM) of the Milky Way and is appropriate for a 1 GeV CR proton diffusing through a 5 G interstellar magnetic field with corresponding Larmor radius \( r_{L,0} \). In equation \( A2, f_{adv} = 0.5 \) is the fraction of CRs that would be removed from the ISM of a SFG by advection in galactic outflows. Large-scale galactic outflows would be common in distant, young SFGs (see, e.g., Frye et al. 2002; Ajiki et al. 2002), being driven by the confluence of feedback from the concentrated starburst episode arising in galactic cores (see Yu et al. 2020, for details about how these may be driven). For plausible outflow velocities, CR advection timescales would indicate that a substantial fraction of CRs could be removed from the nucleus of a SFG by an outflow wind. Due to the prevalence and strength of outflows in the core of SFGs, we adopt a fiducial value of 50% here (as indicated by the timescales shown in Peretti et al. 2019), however we consider the exact fraction is unsettled and would vary substantially between different galaxies and model configurations – e.g. Owen et al. (2019a) estimated a value of \( f_{adv} \sim 0.1 \), but calculated this as a fraction of CRs lost from the entire host galaxy, not just the starburst core. \( V_S \) is the volume of each of the \( N \) sources (if chosen physically, this would correspond to a characteristic size of a SN remnant), and \( A \) is an CR attenuation factor due to their interaction losses within the interstellar medium of the host galaxy. This is written as

\[ A(\gamma_p, r) = \exp \left\{ -\int_{r}^{r_i} g(\gamma_p, r') \, d\gamma_p \right\}, \]  

(A4)

which quantifies the level of attenuation experienced by a beam of CR protons between a source at location \( r \), and some general location \( r' \) at a distance of \( r_i = |r_i - r| \) apart. Here, \( g(\gamma_p, r') = n_H(\gamma_p) \sigma_{\text{pH}}(\gamma_p) \), and \( \sigma \) is the ratio of the free-streaming (i.e. the attenuation length due to CR interactions \( \ell_{\text{pH}} \sim 1/\alpha^2 \) in a non-magnetised medium) and diffusive path lengths of the CRs,

\[ g(\gamma_p, r') = \frac{\ell_{\text{pH}}}{\ell_{\text{diff}}} \left( \frac{c}{4 \, D(\gamma_p) \, n_H(\gamma_p) \sigma_{\text{pH}}(\gamma_p)} \right)^{1/2}, \]  

(A5)

(Owen et al. 2018). The term \( Q_p \) in equation A2 quantifies the CR injection rate discretised by source, such that \( Q_p \, d\gamma_p \) is the rate of injection of CRs within an energy interval \( d\gamma_p \). We define this as

\[ Q_p(\gamma_p, r_i) = S_N(r) \, \frac{L_0}{N} \frac{\partial}{\partial \gamma_p} \left( \frac{\gamma_p}{\gamma_{p,0}} \right)^{-1} \bigg|_{r_i}, \]  

(A6)

where the volumetric CR injection rate is \( S_N(r) \), and the normalisation

\[ L_0 = \frac{L_{\text{CR},\text{eff}}(1 - \Gamma) E_0^{-1}}{E_{\text{max}} - E_1^{-1}}, \]  

(A7)

follows from the CR energy budget set by equation A1. Here, we use a reference energy \( E_0 = 0.1 M_\odot c^2 \) of 1 GeV, and set a maximum CR energy of \( E_{\text{max}} = 50 \) GeV (Peretti et al. 2019). We relax this choice in section 4.2.2, where we consider the impact of different choices of maximum CR energy.

This aim of this study is to assess the contribution of galaxies to the \( \gamma \)-ray background. As such, a detailed formulation of the sub-galactic variations in CR density is not required. Instead, we set \( S_N = 1 \) and so approximate the CR proton density \( n_p \) to be uniform throughout each galaxy interior (correspondingly, we also consider a uniform gas density, of mean value \( n_\text{H} = \langle n_H \rangle \sim 1 \) cm\(^{-3}\), and leave more detailed considerations of the impact of the ISM density and
structure to the future work). This also removes the need for discretisation of the model, so we set \( N = 1 \) in equation A2, which reduces to

\[
\frac{\partial n_p}{\partial r_p} = \frac{R^2 f_{d0b} \mathcal{L}_0}{3(\gamma')^2 f_{d0b}} \frac{\partial}{\partial r_p} \left( \frac{r_p}{\gamma_p,0} \right)^{-1} d\gamma_p, \\
= \frac{35 R^2 f_{d0b} \mathcal{L}_0 \mathcal{A}(\xi_p,\eta) f_{d0b}}{108 D \mathcal{L}_0} \frac{\partial}{\partial r_p} \left( \frac{r_p}{\gamma_p,0} \right)^{-1} d\gamma_p. \quad \text{(A8)}
\]

Here, \( \mathcal{A}(\xi_p,\eta) \) is the mean attenuation of protons as they propagate through the host galaxy (the form of the mean attenuation function \( \mathcal{A}(\cdot) \) for a uniform sphere is given by equation C5), and \((\gamma') = 36R/35 \) is the mean separation length within a spherical volume to radial points uniformly distributed throughout the volume, where \( R \) sets the characteristic size of the volume – in this case, the radius of a characteristic star-forming region of the host galaxy. We also define \( \xi_p = (\rho R/\xi_p)^{1/2}, \) and \( \eta = (\sigma_p \rho n_1)^{-1} \) is the mean free path of protons undergoing hadronic interactions as they propagate through a uniform ISM of density \( n_1 \).

### A1 CR diffusion parameter in star-forming galaxies

Equation A3 sets \( D \) as the diffusion coefficient for CRs in the ISM of their host galaxy. It is specified by both the CR energy, \( \gamma_p \) and the ambient mean magnetic field strength, \( \langle |B| \rangle = |B| \). For such an environment, we set the normalisation value as \( D_0 = 3.0 \times 10^{28} \text{ cm}^2 \text{ s}^{-1} \), which is based on empirical measurements of the diffusion of CRs in the ISM of the Milky Way. It would be appropriate for a 1 GeV CR proton diffusing through a 3\( \mu \)G interstellar magnetic field with corresponding Larmor radius \( r_L \). We argue that there is no strong physical basis to motivate different values in star-forming galaxies, where the processes which set this value are not likely to be different to those seen in the local Universe, and consider that alternative values would not imply more correct physics. \( \xi \) is introduced in equation A3 to encode the interstellar magnetic turbulence. For this, we adopt a value of 1/2 (e.g. Berezinskii et al. 1990; Strong et al. 2007), which is appropriate for a Kraichnan-type turbulence spectrum following a power law of the form \( P_k(k) \sim k^{-3} \), and is thought to be a reasonable description for the turbulence in an ISM (see Yan & Lazarian 2004; Strong et al. 2007).

The magnetic field of the host galaxy is also thought to be driven by star-formation via a turbulent dynamo mechanism during a starburst phase (see e.g. Beck et al. 2012; Latif et al. 2013; Schober et al. 2013), and observational studies favour the rapid development of magnetic fields in protogalaxies, reaching strengths comparable to the Milky Way within a few Myr of their formation (Bernert et al. 2008; Beck et al. 2012; Hammond et al. 2012; Rieder & Teyssier 2016; Sur et al. 2018). The saturation level of such a mechanism, for example introduced by Schober et al. (2013), may be approximated by invoking equipartition with the turbulent energy in the turbulent gas, \( B_{1,\text{sat}} = \left[ 4 \pi \rho/\mu \right]^{1/2} \left[ v_f, f \right] \) where \( \mu = n_p n_1 \) is the local gas density and \( v_f \) is the fluctuation velocity \((v_f \approx R_{\text{gal}}(2n_pG/3)^{1/2}) \) for the protogalaxy, if adopting a pressure with gravity equilibrium approximation – see Schober et al. (2013).

Here, \( f_1 \) represents the deviation from exact equipartition to account for the efficiency of energy transfer from the turbulent kinetic energy to magnetic energy, which simulation work estimates to be around 10% (see, e.g. Federrath et al. 2011; Schober et al. 2013).

### APPENDIX B: ATTENUATION OF \( \gamma \)-RAYS IN MATTER AND RADIATION FIELDS

Section 2.2.2 considers that \( \gamma \)-ray absorption in SFGs is predominantly attributed to \( \gamma \gamma \) pair production in ambient radiation fields. However, other studies have argued that dense interstellar clouds may also have a role. In Lacki (2012), the ability of \( \gamma \)-rays to ionise dense interstellar clouds was discussed. Beams of hadronic CRs can easily undergo interactions in dense environments, where substantial attenuation may arise, leading to the production of leptons, neutrinos and \( \gamma \)-rays. A similar mechanism invoking CR beam dumping is presented in Vereecken & de Vries (2020). The pp inelastic cross section (equation 4) is of order \( \sigma_{\gamma p} \approx 10^{-26} \text{ cm}^2 \), being only weakly dependent on CR energy. The corresponding CR path length due to pp losses in the dense core of a molecular cloud with volume density of \( n_1 \approx 10^5 \text{ cm}^{-3} \) would be \( \ell_p \approx (\sim n_1 \sigma_{\gamma p})^{-1} \approx 300 \) pc. This is substantially larger than the size of a dense core (which would typically extend for less than a pc – see, e.g. Bergin & Tafalla 2007), suggesting CR attenuation in such an environment would be relatively unimportant. Nevertheless, \( \gamma \)-ray production would still arise, with a fraction of around 0.1% of the CR beam intensity undergoing hadronic interactions through the dense core of a molecular cloud. Around 1/3rd of the attenuated CR energy would then be passed to \( \gamma \)-rays, as can be determined from the branching ratios of the intermediate pions (Dermer & Menon 2009). These \( \gamma \)-rays could be attenuated by Bethe-Heitler \( \gamma Z \) pair production \((\gamma + Z \rightarrow e^- + e^+ + Z) \), where \( Z \) is an atomic nucleus – see Lacki (2012; Vereecken & de Vries 2020). In inner regions of SFGs, dense clouds are common: high volume filling fractions of clouds in star-forming nuclei have been inferred for nearby luminous infrared galaxies such as Arp 299 (Sliwa et al. 2012), and inner gas volume densities in some regions of such systems could exceed \( 10^5 \text{ cm}^{-3} \) (e.g. Imanishi et al. 2019). Taking an extreme interpretation of these findings to estimate an upper limit for the attenuative effect of dense clouds on \( \gamma \)-rays in a SFG nucleus, we consider a uniform density medium of \( 10^5 \text{ cm}^{-3} \) throughout a \( R = 0.1 \) kpc nuclear starburst region. At 1 GeV, the Bethe-Heitler \( \gamma Z \) pair production cross section is \( \sigma_{\gamma Z} \approx 10^{-26} \text{ cm}^2 \) (Berestetskii et al. 1980), and is not strongly energy-dependent. This gives a \( \gamma \)-ray attenuation fraction of around 30%, with a mean path-length of \( \ell_{\gamma Z} \approx 0.3 \) kpc. Comparing with Figure 1 (upper panel), this process would dominate over losses due \( \gamma \gamma \) interactions up to around 10 GeV. However, in the lower panel, it can be seen that a 30% reduction in CR flux below 10 GeV would not be of great consequence to the emitted spectrum. As such, we do not consider \( \gamma Z \) losses in SFG nuclei in this work, and leave this to more detailed future studies.

### APPENDIX C: INTERNAL ATTENUATION OF \( \gamma \)-RAYS IN SFGS

The attenuation of \( \gamma \)-rays within a SFG nucleus can be characterised as an average mean attenuation through a homogeneous, isotropic spherical volume. In this scenario, the distance of the path between two points with positions defined by the vectors \( p_1 \) and \( p_2 \) is given by

\[
s^2(p_1, p_2) = r_1^2 + r_2^2 - 2r_1 r_2 \cos \psi \quad \text{(C1)}
\]

where \( \cos \psi = \mu_1 \mu_2 + (1 - \mu_1^2)^{1/2} (1 - \mu_2^2)^{1/2} \cos(\phi_1 - \phi_2) \) are their angular separation, and \( \mu_1 \) and \( \mu_2 \) are the cosines of their angular positions, \( \theta_1 \) and \( \theta_2 \) respectively. The attenuation of a \( \gamma \)-ray beam along each of the separation paths then follows as \( A(p_1, p_2) = \int_0^L \frac{\exp(-\sigma s)}{s} ds \)
exp(–s(p_i, p_j)/ℓ) which, for an ensemble of points p_i following a random uniform distribution, may be written as

\[ \langle A(q) \rangle = \frac{\int_{V_i} A(q, p_i) dV_i}{\int_{V_i} dV_i} \]

\[ = \frac{3}{4\pi R^3} \int_{4\pi} d\Omega_i \int_{r_i=0}^R \exp(-s(q, p_i)/\ell) r_i^2 dr_i , \] (C2)

where dV_i = r_i^2 dr_i dΩ_i is the differential element corresponding to the volume V_i occupied by the points distribution p_i in the continuous limit. Under spherical symmetry, the mean attenuation between some reference position q at a radius r within the sphere to the points in the ensemble p_i (again, in the continuous limit) follows by averaging over S_r, the spherical surface specified by r, i.e:

\[ \langle A(r) \rangle = \frac{\int_{S_r} A(q) dS_r}{\int_{S_r} dS_r} \]

\[ = \frac{2}{8\pi R^3} \int_{\Omega} d\Omega \frac{\partial}{\partial r} \int_{r'=0}^R (s(q)) r' dr' \] (C3)

where the surface element is separated into its radial and angular components, for which dS_r = 2\pi r' dr' dΩ. This may be discretised and reduced into a numerical Monte-Carlo problem by distributing a uniform random ensemble of N = 10^6 points within a sphere of radius R to calculate the mean attenuation \( \langle A(r) \rangle \) between the points and a radial line r from the origin r = 0 to the edge of the sphere r = R. The total characteristic attenuation through the sphere then follows by taking the mean value of A(r) along the radial line from r = 0 to R, i.e:

\[ A(R, \ell) = \frac{1}{R} \int_{r=0}^R A(r) dr . \] (C4)

The resulting characteristic attenuation is then specified by only the size of the sphere (describing the extent of the nuclear core of a star-forming galaxy) and the effective mean path length of the γ-rays within the sphere’s internal medium (density and radiation field), L_{mfp}. Equation C4 may be well-approximated by a Gaussian function,

\[ A(\zeta) = \exp(-\zeta^2) , \] (C5)

where \( \zeta = (R/L_{mfp})^{1/2} \). We find this approximation gives an error of less than 1% for all reasonable values of \( \zeta \).

APPENDIX D: EGB CLUSTERING AND POISSON NOISE

The SFG power spectrum \( P_\ell(k, z) \) would imprint a signature in the EGB, and this could be measured from γ-ray background observations using the auto-correlation function (e.g. Peebles 1980; Inoue et al. 2013b),

\[ \langle C(\theta) \rangle = \langle \delta I(l_1) \delta I(l_2) \rangle \]

\[ = \frac{1}{16\ell^2} \int_0^{\infty} dl_1 \int_0^{\infty} dl_2 \xi(l_1 - l_2) I(l_1) I(l_2) , \] (D1)

where \( \xi(\ldots) \) is the two-point correlation function, and I(l) is the intensity of the EGB at some position specified by the co-moving vector l, of (co-moving) length l. \( \delta I(l) \) is the intensity fluctuation, being the deviation of I at some position l from its mean value, and \( \theta \) is the angular separation of the positions l_1 and l_2. In general, the clustering (or correlation) term of the angular power spectrum can be defined as the Fourier Transform of the auto-correlation function.

In the case of EGB analyses, we consider a γ-ray signal in a 2-dimensional space \( C^F \), which can be split into two components,

\[ C^F = C^P + C^C \] (D2)

where \( C^P \) is an isotropic Poisson noise term (an auto-correlation term), and \( C^C \) is the clustering term between points of angular separation \( \theta > 0 \). Previous work assessing the AGN contribution to the EGB (e.g. Inoue et al. 2013b) found the Poisson term to be comparable to the clustering term. In this case, the EGB contribution was comprised of a population of bright, unresolved point sources. The SFG contribution would be different, with the γ-ray background emission being dominated by a much larger number of galaxies, each being fainter than a typical AGN source. As such, the \( C^P \) term would presumably be much smaller than the \( C^C \) term here.

We compute the Poisson and correlation terms separately. For the clustering term, we only require (isotropic) angular separations such that it reduces to

\[ C^C = \int_{\theta>0} d^2 \theta e^{-i \theta} C(\theta) , \] (D3)

where the separation angle \( \theta \) is non-zero. Assuming that the mean signal intensity is the same between the positions at l_1 and l_2 and that the signal I is statistically isotropic and homogeneous both in space and its projection onto the sphere, and if adopting the notation \( r_2 = l_2 - l_1 \) and \( r_1 = (l_2 + l_1)/2 \) for convenience (cf. the Limber 1953 approximation), we arrive at

\[ C^C = \frac{1}{16\pi} \int_{\theta>0} d^2 \theta e^{-i \theta} \int_0^{\infty} dr_1 \int_{-r_1}^{r_1} dr_2 \xi(r_2 \hat{r} + r_1 \hat{\theta}) I^2 \]

\[ - \frac{1}{16\pi^2} \int_{\theta>0} d^2 \theta e^{-i \theta} \int_0^{\max} d^2 V_c \int_0^{\infty} dz \frac{\xi(r_2 \hat{r} + r_1 \hat{\theta}) L^2_\gamma(z)}{r_2^2 (1 + z)^2} , \] (D4)

where we approximate the limits of \( r_2 \) to be ±∞. Here, \( r_p \) is the proper distance in the \( \hat{r} \) direction, \( V_c \) is a comoving volume and \( \hat{r} \) and \( \hat{\theta} \) are introduced as unit vectors in the radial direction towards the background (thus being a function of redshift) and the direction between the two points at l_1 and l_2, respectively. \( L_\gamma(z) \) is the γ-ray luminosity of the source population at a distance of redshift \( z \).

The two-point correlation function is related to the underlying power spectrum of the γ-ray source population by a Fourier Transform,

\[ \xi(r_2 \hat{r} + r_1 \hat{\theta}) = \frac{1}{8\pi^3} \int d^3 k P_\ell(k) \exp \left( ik \cdot [r_2 \hat{r} + r_1 \hat{\theta}] \right) . \] (D5)

Here, \( k \) is a proper wavenumber, defined as the magnitude of the wavevector \( k \). We may combine and decompose the exponential terms of equation D4 into

\[ \exp \left( ik \cdot [r_2 \hat{r} + r_1 \hat{\theta}] \right) - i(\hat{r} \cdot \theta) = \exp \left( ir_2 k_|| \cdot \hat{r} + i(\hat{r} \cdot [r_1 k_\perp - l]) \right) \] (D6)

if expressing \( k \) in terms of parallel and perpendicular components to the direction \( \hat{r} \). The integrals over \( r_2 \) and \( \theta \) in equation D4 may then be readily evaluated as:

\[ \int_{\theta>0} \int_{r_2=0}^{\infty} d^2 \theta dr_2 \exp \left( ir_2 k_|| \cdot \hat{r} + i(\hat{r} \cdot [r_1 k_\perp - l]) \right) \]

\[ = \frac{8\pi^3}{r_2^2} \delta(k_{||,p}) d^2 (r k_{\perp, p} - l_p [1 + z]) , \] (D7)
where the second step converts comoving coordinates to proper coordinates such that \( r_{p,2} = (1 + z) r_2 \), and subscript \( p \) denotes proper coordinates (as used in the wavenumbers), and the equivalence of the \( k_\perp \) components follows from the earlier assumption that the signal is isotropic in the sky plane. Equation D4 then reduces to

\[
C^\gamma_L = \frac{1}{16\pi^2} \int_0^{z_{\text{max}}} \frac{d^2V_c}{dz d\Omega} \frac{L_c^2(z)}{r_p^2 (1 + z)^2} \times \int dk_\parallel d^2k_\perp P(k_\parallel) \delta(k_\parallel) \frac{\delta^2(r_p k_\perp - l_p [1 + z])}{r_p^2} \frac{1}{16\pi^2} \int_0^{z_{\text{max}}} \frac{d^2V_c}{dz d\Omega} L_c^2(z) P(l_p [1 + z]/r_p) \frac{1}{r_p^2 (1 + z)^2}, \tag{D8}
\]

after integrating over the orthogonal wave-vector components, which is the combined contribution of SFGs to the EGB up to some redshift \( z_{\text{max}} \). In differential units of flux, this gives

\[
C^\gamma_L(E_\gamma) = \int_0^{z_{\text{max}}} \frac{d^2V_c}{dz d\Omega} \frac{1}{4\pi D_L} P \left( \frac{l_p}{r_p} [1 + z] \right) \left( \frac{dF_\gamma(E_\gamma, z)}{dE_\gamma} \right)^2 \tag{D9}
\]

which is equation 28 in the main text. Here,

\[
\frac{dF_\gamma(E_\gamma, z)}{dE_\gamma} = \frac{1 + z}{4\pi D_L^2} \frac{dL_c^\gamma(E_\gamma [1 + z])}{dE_\gamma}, \tag{D10}
\]

which describes the redshift-dependent emission of \( \gamma \)-rays from the source population of SFGs, thus incorporating the internal and external \( \gamma \)-ray attenuation/reprocessing models, and the co-moving number density of SFGs. \( D_L \) is the luminosity distance, defined as

\[
D_L = (1 + z) \left( \frac{c}{H_0} \right) \int_0^z \frac{dz'}{E(z')}, \tag{D11}
\]

for \( E(z) = \left[ \Omega_{m,0}(1 + z)^3 + \Omega_\Lambda,0 + \Omega_k,0(1 + z)^2 \right]^{1/2} \), where terms retain their earlier definitions. The Poisson term \( C^\gamma_L \) is then given by

\[
C^\gamma_L(E_\gamma) = \int_0^{z_{\text{max}}} \frac{d^2V_c}{dz d\Omega} \left( \frac{dF_\gamma(E_\gamma, z)}{dE_\gamma} \right)^2 \tag{D12}
\]

which is equation 29 in the main text.

**APPENDIX E: COMPUTATIONAL METHOD**

We directly compute the EGB anisotropy statistic at \( z = 0 \) for an energy \( E_\gamma \) using equation 28. To do this, we adopt a numerical approach where the volume to redshift \( z_{\text{max}} \) containing the EGB source population is discretised into \( N_\gamma \) shells. The contribution from each shell to the EGB at \( z = 0 \) is calculated by solving equation 20 subject to the boundary condition set by the \( \gamma \)-ray intensity at the originating shell. This is the combined contribution of the (primary) SFG \( \gamma \)-ray emission from the shell, taken through the volume between shell \( i \) and \( i + 1 \), plus a (secondary) background contribution to that shell – i.e. the propagated emission from higher redshift shells and their cascaded component reprocessed to an energy \( E_\gamma \) (equation 25).

The primary \( \gamma \)-ray emission of the SFG galaxy population at each shell is found by integrating the \( \gamma \)-ray contribution of a galaxy per star-formation rate (computed from the SFRF – see section 3 for details) between 1 and 10000 M_{\odot} yr^{-1}, using a logarithmic trapezium-rule with \( N_{\text{SFR}} \) steps. The original \( \gamma \)-ray spectral emissivity per galaxy (equation 6) is computed in the same manner, using \( N_E \) steps. The primary \( \gamma \)-ray contribution is convolved with the SFG power spectrum to encode the spatial dependence of the emission.

The secondary \( \gamma \)-ray emission function in the transfer equation 25 depends on the electron injection rate, given by equation 24. This, in turn, is set by the absorption of \( \gamma \)-rays propagating to a shell \( i \) from more distant shells. The integral in equation 24 must therefore be evaluated for each step, \( i \). This is achieved by a second level of discretisation, with a further redshift grid defined from \( j = 1 \) to \( N_{\gamma,2} \), with the primary \( \gamma \)-ray intensity computed at each ‘sub’-shell (again, using equation 20). At each sub-shell, the contribution from the full spectra of higher-redshift \( \gamma \)-rays must be considered, as a fraction of the spectrum at each energy will provide a contribution to lower energy \( \gamma \)-rays, via the cascade process. The double integral of 25 is computed by discretisation of the dimensionless variable \( x_\gamma \) into \( N_\gamma \) steps (for the inner integral), and the outer integral follows likewise with \( N_\gamma \) steps. We found a simple trapezium-rule numerical integration method to be sufficient for both of these.

The inner redshift integral could be computed by a simple adaptive Runge-Kutta (RK) Fehligbargh 4th order scheme (Press et al. 1992), while we required the greater numerical stability afforded by an implicit RK 4/5 scheme for the outer redshift grid. For this, we used the RADAU5 solver of Hairer et al. (1993). Strictly, a redshift grid with a finer resolution than the typical absorption length of \( \gamma \)-rays in the EBL should be adopted. However, we found that sub-grid variations in the \( \gamma \)-ray intensity due to this attenuation and cascade re-emission made little difference to our results. As such, a coarser grid could be safely adopted, and a choice of \( N_\gamma = 100 \) and \( N_{\gamma,2} = 100 \) was found to give results which varied by less than 1% compared to higher resolution grids (we compared to grid resolutions increased by a factor of two in both cases). To achieve a comparable 1% level of numerical accuracy, we found minimum resolutions of \( N_{\gamma,1} = 10 \) and \( N_\gamma = 10 \) were sufficient to numerically evaluate the integrals in equation 25. Higher resolution discretisation was found to be necessary for the source function \( \gamma \)-ray emission computations (sections 2.1.2 and 3.1, respectively), with both \( N_{\text{SFR}} = 50 \) and \( N_E = 50 \).