Giant cosmic ray halos around M31 and the Milky Way

Stefano Gabici, a,∗ Sarah Recchia, b, c Felix Aharonian d, e and Viviana Niro a

a Université de Paris, CNRS, Astroparticule et Cosmologie, F-75006 Paris, France
b Dipartimento di Fisica, Università di Torino, via P. Giuria 1, 10125 Torino, Italy
c Istituto Nazionale di Fisica Nucleare, Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
d Dublin Institute for Advanced Studies, 31 Fitzwilliam Place, Dublin 2, Ireland
e Max-Planck-Institut für Kernphysik, Postfach 103980, D-69029 Heidelberg, Germany

E-mail: gabici@apc.in2p3.fr, sarah.recchia@unito.it

Recently, a diffuse emission of 1-100 GeV γ-rays has been detected from the direction of M31, extending up to 200 kpc from its center. The interpretation of the extended γ-ray emission by the escape of cosmic rays produced in the galactic disk or in the galactic center is problematic. In this paper, we argue that a cosmic ray origin (either leptonic or hadronic) of the γ-ray emission is possible in the framework of non standard cosmic ray propagation scenarios or is caused by in situ particle acceleration in the galaxy’s halo. Correspondingly, the halo is powered by the galaxy’s nuclear activity or by the accretion of intergalactic gas. If the formation of cosmic ray halos around galaxies is a common phenomenon, the interactions of cosmic ray protons and nuclei with the circumgalactic gas surrounding Milky Way could be responsible for the isotropic diffuse flux of neutrinos observed by Icecube.
1. Introduction

In the last few years the Andromeda galaxy (M31) has been detected in γ-rays Fermi-LAT [1, 2], with an integrated γ-ray luminosity above 100 MeV of $\sim 6.6 \times 10^{41}$ s$^{-1}$, less than a factor of 2 to that of the MW [1]. Contrary to the MW, where the emission correlates spatially with the gaseous disk, for M31 it appears to be concentrated within the inner $\sim 5$ kpc region [2], while recently the possible existence of Fermi Bubbles-like structure emanated from the central region of M31 has also been reported [3].

A recent analysis of Fermi-LAT data found the presence of an extended γ-ray emission from a very large halo surrounding M31 [4]. The authors studied a region of $28^\circ \times 28^\circ$, that corresponds to a projected radius of $\sim 200$ kpc from the center of M31 and found an excess that extends up to $\sim 120 – 200$ kpc. The authors performed an accurate modeling of the MW foreground emission and added to the analysis a spherically symmetric template, centered on M31 and divided into three regions: the inner galaxy (IG), a $\sim 5.5$ kpc radius region that contains the emission from the inner galaxy [2], the spherical halo (SH), an intermediate ring which extends up to $\sim 120$ kpc and the outer halo (OH), a ring of $\sim 120 – 200$ kpc, from which they extracted the γ-ray spectra. Since the northern part of these regions, in particular that of the OH, partially overlap with the MW disk, the authors reported also the spectra for the southern and northern part of the SH and OH, and concluded that, while the excess from the SH region is likely associated to the halo of M31, the origin of that from the OH is much less clear and may be rather related to the MW, or to something else. A spectral fit with power-law plus exponential cut-off parametrization and with a pure power-law is reported by [4]. The former reads

$$I_{SH} \approx 9.8 \times 10^{-11} E_{GeV}^{-1.9} e^{-E_{GeV}/11.6\text{MeV}} \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$$

where $E_{GeV}$ is the photon energy in GeV, while the latter is shown in Fig. 1. The total γ-ray luminosity for the northern plus southern part of SH equal to:

$$L_\gamma \approx 1.7 – 1.9 \times 10^{39} \text{erg/s}.$$  

where the lower(upper) value corresponds to the power law plus cutoff(pure power law) fit to data.

Based on the large extension of γ-ray halo, on the intensity and spectrum of the various components, the authors of [4] suggest a dark matter interpretation for such emission [5] and conclude that it is unlikely that a major fraction of the emission is due to cosmic ray (CR) in the halo of M31.

Here we show, instead, that a CR origin of such extended emission is feasible and also quite natural. It would require a non-standard mechanism of CR propagation into galactic halos or the acceleration of particles directly in the halo, able to produce a giant CR halo of radius $\sim 100$ kpc around M31. Interestingly, the possible presence of such large halos was proposed for both the MW and M31 [6, 7], and it was shown that such a halo around the MW could explain the diffuse Icecube neutrino flux [7], and a subdominant fraction of the isotropic gamma-ray background [6].

2. Hadronic and leptonic origin of the γ-ray emission: energetics

In the leptonic scenario the γ-ray emission from the SH is the result of inverse Compton scattering (ICS) of energetic electrons off CMB photons. The galactic magnetic field at such
distances from the disk is likely below $\sim 1 \mu G$ so that synchrotron losses are probably negligible in the SH. Thus, the SH emission, detected in the range $\sim 1 - 100$ GeV, would be produced by electrons of energy $\sim 0.6 - 6$ TeV, and the cutoff at $E_{\gamma} = 11.6$ GeV would correspond to electrons of energy $E_{\text{max},e} \approx 1.9$ TeV. The electron energy loss time in the CMB field is given by (see [8] and references therein)

$$\tau_{\text{CMB}} \sim 1.3 \times 10^6 E_{\text{TeV}}^{-1} \text{ yr.}$$

For any realistic interstellar diffusion, such short timescale implies that such particles cannot likely come from the galactic disk/center of M31 and have to be accelerated directly in the SH. The minimal energy requirement is obtained in the assumption that the ICS timescale is shorter than any other relevant time-scale so that the $\gamma$–ray production takes place in a calorimetric regime, namely the observed $\gamma$-ray luminosity (Eq. 2) equals the electron luminosity:

$$L_e = L_\gamma \approx 1.7 - 1.9 \times 10^{39} \text{ erg/s.}$$

This estimate refers only to electrons in the energy band $\approx 0.6 - 6$ TeV, and is of the same order than the total estimated power of CR electrons in the disk of the MW [9].

In the hadronic scenario $\gamma$–rays are produced in proton-proton interactions between CRs and the ambient gas. Photons of energy $E_\gamma$ are produced by CR protons of $E_p \approx 10 E_\gamma$, and the required CR proton luminosity $L_p$ is related to the $\gamma$-ray luminosity $L_\gamma$ by [7]:

$$L_p = 3 L_{\gamma}/f, \quad f = 1 - e^{-\tau_{\text{res}}/\tau_{pp}},$$

where $\tau_{\text{res}}$ is the residence time of CRs in the halo and $\tau_{pp}$ is the timescale for pp interactions,

$$\tau_{pp} \sim 7.1 \times 10^{10} n_{H,-3}^{-1} \text{ yr}$$

where $n_{H,-3}$ is the halo hydrogen density in units of $10^{-3}$ cm$^{-3}$. Typical values of $n_H$ expected at $\sim 100$ kpc from the disk of the MW are $\sim 10^{-4} - 10^{-3}$ cm$^{-3}$ (see [8] and references therein), for which the loss time is larger than the age of the Universe. Taking $\tau_{\text{res}} = 10^9 \tau_{\text{res},9} \text{ yr}$ the CR proton luminosity needed to explain the $\gamma$-ray observations of the SH in M31:

$$L_p \approx 1.8 \times 10^{41} \tau_{\text{res},9}^{-1} n_{H,-3}^{-1} \text{ erg/s.}$$

Notice that for $\tau_{\text{res},9} \gtrsim 1$ and $n_{H,-3} \sim 1$, such luminosity is of the same order of that invoked to explain the population of CR protons observed in the disk of the MW [9].

3. Origin of the radiating particles

In the typical scenario for the origin of Galactic CRs, they are accelerated in sources located in the disk and escape form the Galaxy as due to diffusion (produced by the CR scattering on the turbulent interstellar magnetic field) and advection in a Galactic wind[9]. A similar situation is likely present also in M31 and one may think that the SH is the result of particles transport from the disk/GC of M31. However such scenario cannot work. Indeed, electrons would lose energy before reaching distances of few tens kpc from the disk, while any typical model of CR proton
advective/diffusive propagation from the disk or from the galactic center (GC) inevitably leads to a substantial decrease of the CR density with the distance from the disk. In this case, fitting the $\gamma$-ray flux observed from the SH would imply a very large intensity of CR protons in the disk of M31, that should be observed as a bright $\gamma$-ray source, which is not observed (see [8] for a detailed discussion).

Motivated by the detection of the Fermi Bubbles in the MW [10], which are likely the result of past nuclear activity in the CG of the MW, and by the possible existence of such structures also in M31 [3], a possible origin of the SH emission may be CR protons produced in periodic episodes of nuclear activity in M31 and transported into the SH within buoyant bubbles. If $\nu_B = 10^{-2} \nu_{B,-2} \text{Myr}^{-1}$ is the frequency of the episodes of nuclear activity in M31, the effective rate at which CR protons may be injected into the halo is $L_p = \eta E_B \nu_B \sim 3.2 \times 10^{41} \eta E_B \nu_{B,57} \nu_{B,-2} \text{erg/s}$, where $E_{B,57} = E_B / 10^{57} \text{erg}$ [11, 12] and $\eta$ is a CR conversion efficiency that also accounts for possible adiabatic energy losses during the transport to the halo. Comparing this expression with the energy requirement in Eq. 7 one finds $\eta \approx 0.56 \tau_{\text{es},9}^{-1} n_{H,-2}^{-1} E_{B,57}^{-1} \nu_{B,-2}$.

CRs can then be transported away from the disk into the galactic halo by CR-inflated-buoyant-bubbles before their disruption by plasma instabilities, such as to avoid a an overproduction of $\gamma$-rays in the disk, contrary to the case in which CRs diffuse/advect directly from the galactic disk or the GC. Such buoyant bubbles, often found in the central regions of clusters of galaxies [13, 14], have been observed also in galaxies [15]. Bubbles are typically found to rise with a velocity of the order of a fraction of the sound speed ($\approx 100 \text{ km/s}$ in the hot diffuse circumgalactic gas [16]), while their typical lifetime could be longer $\sim 10^9 \text{ yrs}$ due to the stabilizing action of a magnetic field (see [16] and references therein). Once released in SH, CRs will spread diffusively to fill a region of $\approx 100 \text{ kpc}$ radius, with a diffusion coefficient which typically increases with the particle energy. When the contributions of different bubbles overlap a quasi-stationary CR population in the SH while in the IG, a possibly intermittent Fermi Bubble like structure may be observed. The former case corresponds to particle energies such that the diffusion time $\tau_{\text{es},9} \sim R_{\text{SH}}^2 / (6 D)$ over a region of size $R_{\text{SH}} \sim 100 \text{ kpc}$ is shorter than the typical time between episodes of nuclear activity $1 / \nu_B$. At large enough energy, this condition is no longer valid and protons will populate the halo intermittently. This would reflect in a spectral cut-off at the particle energy, $E_p$, such that $D(E_p) = 5 \times 10^{30} R_{100}^2 \nu_{B,-2} \text{ cm}^2 / \text{s}$, and could explain the power law plus cut-off fit to $\gamma$-ray data (black points in Fig. 1), provided that $E_p \approx 100 \text{ GeV}$ (see [8] and references therein). If the diffusion coefficient is rather flat in energy, and/or if all CRs injected during the entire lifetime of M31 are confined in the SH, no spectral cut-off is expected, a case consistent with the pure power law fit to data (see Fig. 1).

In a lepton scenario, in situ acceleration is necessary. The maximum energy $E_{\text{max},e}$ of accelerated electrons at a spherical shock of radius $R_s$ can be estimated by equating the acceleration time scale to the energy loss time $\tau_{\text{loss}}$ as due to ICS in the assumption that particle acceleration at shocks is accompanied by an amplification of the magnetic field, and reads (see [8] for the details) $E_{\text{max},e} \approx 24 n_s^{1/4} [u_s h_{1,3}^{-1/4} (\xi_B / 0.035)^{1/4}]$ TeV, where $n_0 = 10^{-4} n_{0,-4} \text{ cm}^{-3}$ is the number density of the intergalactic gas, $u_s = 10^3 u_{s,3} \text{ km/s}$ is the velocity at which matter flows into the shock, $\xi_B = 3.5\%$ of the shock ram pressure is converted into magnetic field energy. Given that the observed $\gamma$-ray
spectrum extends up to at least a photon energy \( E_\gamma \approx 10 \) GeV, the electron spectrum should extend at least up to an energy \( E_{\text{max,e}} \approx 2 \) TeV, which requires \( u_{s,3} \geq 0.43 n_{0,-4}^{-1/12} \left( \frac{E_\gamma}{63 \text{ GeV}} \right)^{-1/12} \). This velocity is remarkably close to the free fall velocity at the edge of the SH \( v_f \sim 0.29 \times 10^3 M_{12}^{1/2} R_{SH,2}^{-1/2} \) km/s, where \( M = 10^{12} M_{12} M_{\odot} \) is the total mass of M31. Moreover, note that the radius of the SH, \( R_{SH} = 100 R_{SH,2} \) kpc, is of the same order of the virial radius of the system. Finally, the energy flow across the shock would be approximately:

\[
L_s \approx \left( 4\pi R_{SH}^2 \right) \frac{Q_0 v_f^3}{2} \sim 3.4 \times 10^{42} R_{SH,2}^{1/2} n_{0,-4} M_{12}^{1/2} \text{ erg/s} \tag{8}
\]

which could accommodate the energy requirement given by Eq. 4. Thus, it is possible to speculate that particles are accelerated at a spherical accretion shock. However it is important to remark that the existence of accretion shocks around galaxies is matter of debate (see [8] and references therein).

The shock may also be formed as a result of the GC nuclear activity. In this case an estimate of the shock velocity \( v_s \) for a shock radius \( R_s \sim R_{SH} \) and a overall duration of the nuclear activity of \( \tau_{\text{GC}} = 10^9 \tau_{\text{GC,9}} \) is given by (see [8] and references therein) \( u_s \approx 0.2 \times 10^3 L_{\text{GC,43}}^{1/5} n_{0,-4}^{1-15/2} \) km/s, where \( L_{\text{GC}} = 10^{43} L_{\text{GC,43}} \text{ erg/s} \) is the time averaged rate of energy injection due to galactic nuclear activity. Notice that the reference value for the energy injection rate is much smaller than the Eddington luminosity of the central supermassive black hole of mass \( M_{BH} \). \( L_{Edd} \sim 1.3 \times 10^{46} \left( \frac{M_{BH}}{10^8 M_{\odot}} \right) \) erg/s.

In both cases, the CR electron luminosity needed to explain observations (Eq. 4) can be account for with an acceleration efficiency at such suck of the order of \( \eta \approx 10^{-4} \).

Together with electrons also CR protons can be accelerated at the shock, probably with a much larger efficiency, as inferred from the study of Galactic CRs [9]. The maximum acceleration energy of protons \( E_{\text{max},p} \) can be obtained posing two conditions: \( i) \) by equating the acceleration time \( \tau_{\text{acc}} \) to the age of the shock \( \tau_s = 10^9 \tau_{s,9} \) yr, since the proton-proton loss time is typically larger mechanism (see Eq. 6); \( ii) \) protons must remain confined in the accelerator, namely diffusion length at energy \( E_{\text{max},p} \) should be smaller than the size of the accelerator (\( \sim R_{SH} \)). The two conditions provides similar values of \( E_{\text{max},p} \), so we consider here the former, which gives \( E_{\text{max},p} \approx 4.6 \times 10^2 u_{s,3}^{3/2} \tau_{s,9} n_{0,-4}^{1/2} \left( \frac{E_\gamma}{63 \text{ GeV}} \right)^{1/2} \) PeV. Thus, the acceleration of CR protons can proceed well beyond the \( \sim 10-100 \) TeV needed to explain the \( \gamma \)-ray observations. The latter can be accounted for by an acceleration efficiency at the level of \( \approx 1\% \).

4. A giant CR halo around the MW and the origin of Icecube neutrinos

As for the MW, the existence of a \( \sim 100 - 200 \) kpc CR halo was proposed in [7] as a possible origin of the diffuse flux of multi-TeV neutrinos detected by Icecube. The differential isotropic flux of astrophysical neutrinos (all flavors, neutrinos plus antineutrinos) measured at Earth can be fitted with a power law [17] \( \Phi^C_{\nu}(E_\nu) \sim 6.37 \times 10^{-18} \left( \frac{E_\nu}{100 \text{ TeV}} \right)^{-2.87} \text{ GeV}^{-1}\text{cm}^{-2}\text{s}^{-1}\text{sr}^{-1} \). The integrated isotropic flux is \( F_{\nu}(> 100 \text{ TeV}) \sim 1.2 \times 10^{-10} \text{ erg/cm}^2/\text{sr} \). If such neutrinos are produced in the MW halo at a typical distance of \( R_H = 10^2 R_{H,2} \) kpc, then the differential neutrino emissivity from the entire MW is \( Q^M_{\nu}(E_\nu) = (4\pi)^2 \Phi^C_{\nu}(E_\nu) R_H^2 \) and the related luminosity
Giant cosmic ray halos

Stefano Gabici

Figure 1: left: Expected $\gamma$-ray emission from M31 (shaded cyan region) obtained rescaling the predictions for the halo of the Milky Way. The black data points and the shaded gray region show Fermi-LAT observations of M31 [4] right: Isotropic diffuse neutrino and $\gamma$-ray emission observed by Icecube [17] (black) and Fermi-LAT [18] (blue data points). Solid lines are predictions for the neutrino (green) and $\gamma$-ray (pink) resulting from the interactions of CR protons with ambient gas in a $\sim 100$ kpc halo surrounding the MW.

$L_{\gamma}^{MW}$ ($> 100$ TeV) $\approx 1.8 \times 10^{39} R_{H,2}^2 \text{erg/s}$, and in the assumption of a proton-proton interaction origin, an isotropic diffuse $\gamma$-ray emission above $\sim 100$ TeV is also expected, with a luminosity comparable to that of neutrinos. The isotropic diffuse fluxes of $\gamma$-rays and neutrinos measured by Fermi-LAT [18] and Icecube [17], respectively, are shown in Fig. 1, together with the predictions from the $\gamma$-ray and neutrino emission from CR proton-proton interactions in the MW halo, estimated assuming a proton spectrum $\propto E_p^{-2} \exp(-E_p/20 \text{ PeV})$ normalized in such a way to contain a total energy equal to $W_p \sim 4.6 \times 10^{57} R_{H,2}^2 \text{erg}$ erg. The cutoff in the $\gamma$-ray spectrum reflects the fact that very high energy photons are absorbed due to pair production in the CMB.

Given the similarities between the MW and M31, it seems natural to compare the hadronic scenarios proposed above to explain the neutrino production in the former and the $\gamma$-ray emission in the latter. The observational constraints for the gas density in the halos of the two galaxies give very similar values (see e.g. [19, 20]), while the CR proton density in the halo would scale linearly with the central SMBH mass in a scenario where the CR content is the linked to the galactic nuclear activity or with the total mass of the galaxy to the power $3/2$ in an accretion shock scenario. The ratio between the mass of M31 and of the MW is $\sim 2$, while the ratio between the SMBH masses is $\sim 33$ (see [8] and references therein). Thus, if the CR confinement time is similar for the two galaxies, of the order of gigayears, one may suppose that the $\gamma$-ray and neutrino luminosities of M31 are a factor of $\sim 2.8 \sim 33$ than those of the MW. Such rescaled fluxes are shown in Fig. 1, together with the spectrum of M31 observed by Fermi-LAT for a power law plus exponential cutoff modeling of data (black points and shaded gray region) and for a power law model (blue line and shaded region) [4]. The good agreement our predictions and Fermi-LAT data suggests a possible common origin of the GeV $\gamma$-ray emission from the halo of M31 and the neutrino emission from the halo of the Milky Way, depending on the presence or absence of a cutoff/steepening in the $\gamma$-ray spectrum of M31.
5. Conclusions

Motivated by the recent discovery in Fermi-LAT of a giant γ-ray halo of size $\sim$ 100-200 kpc around M31, and by the existence of a similarly extended gaseous halo around the MW, we investigated the consequences of the possible presence of a giant CR halo surrounding both galaxies.

We concluded that, if CRs can be confined for long times in the galactic halos (order of gigayears), both the isotropic diffuse neutrino emission observed by Icecube (as proposed by [7]), and the extended γ-ray emission measured by Fermi-LAT around M31 could be explained in terms of CR protons interacting with the circumgalactic gas. Such particles may be produced in the GC of M31 and then transported into the halo by means of buoyant bubbles, or may be accelerated \textit{in situ} at a large shock in the SH region. In the former case, the morphology of the emission is expected to be similar to that of Fermi Bubbles, but much more extended, while in the latter case, the emission is expected to be roughly spherically symmetric. The required time averaged luminosity of CRs injected in the halo is of the order of $\gtrsim 10^{40} - 10^{41}$ erg/s, comparable to the estimated luminosity of CR sources in the Galactic disk. Such scenario predicts a multi-TeV γ-ray emission from the halo of M31 that is within the reach of instruments such as LHAASO.

If gaseous halos are a common feature of MW and M31-like galaxies they might all emit both γ-rays and neutrinos, which however are beyond the capabilities of current instruments for distances significantly larger than M31 (see [8] for a discussion).

Alternatively, the M31 γ–ray emission could be of leptonic origin. In this case no neutrinos would be expected from M31.

References


